THE EVOLUTIONARY HISTORY OF MADAGASCAR'S TENRECS (MAMMALIA: TENRECIDAE): SYSTEMATICS, PHYLOGEOGRAPHY, AND SPECIES DELIMITATION

By

Kathryn Michelle Everson, B.S.

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in

Biological Sciences

University of Alaska Fairbanks

August 2018

APPROVED:

Dr. Link Olson, Committee Chair

Dr. Sharon Jansa, Committee Member

Dr. Derek Sikes, Committee Member

Dr. Naoki Takebayashi, Committee Member

Dr. Diane Wagner, Chair

Department of Biology and Wildlife

Dr. Anupma Prakash, Dean

College of Natural Science and Mathematics

Dr. Michael Castellini, Dean

UAF Graduate School

Abstract

Madagascar is renowned for its exceptionally diverse, unique, and threatened biota, yet much of the island's flora and fauna remains undescribed, and the underlying drivers of diversification and endemism are poorly understood. The family Tenrecidae is one of four extant terrestrial mammal lineages to have colonized and diversified on Madagascar from continental Africa. The goal of this dissertation is to elucidate the evolutionary history of tenrecs at both deep and shallow time scales, and to use tenrecs as a proxy to understand the drivers of diversification on Madagascar. In Chapter 1 I generate the first rigorously inferred phylogeny of tenrecs to include every currently recognized species, revealing that they colonized Madagascar 30-56 million years ago. I also demonstrate that speciation rates have been higher in humid habitats compared to arid habitats – a finding that sets the groundwork for my next three chapters.

To better understand the patterns and processes of diversification in the humid forest, I next explore the phylogeography of a species endemic to that region, *Oryzorictes hova*, in Chapter 2. Using genetic and morphometric data, I identify three populations (later identified as cryptic species) within *O. hova* that correspond to northern, central, and southern regions of the island. The same phylogeographic pattern has been observed in some of Madagascar's other humid-forest taxa, and it had been hypothesized that population structure is driven by low-elevation breaks between Madagascar's northern, central, and southern highlands. In Chapter 3, using genetic data from 20 small mammals and five reptiles codistributed along the island's humid-forest belt, I find this structure is directly related to the distribution of high-elevation areas and is congruent (spatially and temporally) across many species. This result demonstrates that the highlands have played an important role in recent diversification on Madagascar, most likely by functioning as refugia during Quaternary climate cycles.

Finally, in Chapter 4 I continue to explore diversification in Madagascar's humid forests by studying species limits and patterns of gene flow in a clade of shrew tenrecs endemic to that region (*M. fotsifotsy, M. soricoides*, and *M. nasoloi*). Using a massively multi-locus genetic dataset, I demonstrate that *M. soricoides* and *M. fotsifotsy* (which are broadly sympatric) have hybridized in the past, and that this has caused conflicting phylogenetic results between different genetic datasets. I also recover two distinct clades of *M. fotsifotsy*: one that occurs only in the far north of Madagascar, and one that is widespread and broadly sympatric with *M. soricoides*. Evidence of reproductive isolation, plus subtle but significant morphometric differences between these clades, lead me to recognize them as distinct species.

While I accomplished my primary aim of clarifying the phylogeny and taxonomy of Madagascar's tenrecs, my findings will also be important to scientists outside that initial scope. This research illuminates one of the mechanisms by which Madagascar's flora and fauna became so diverse – namely, that diversification has been driven by latitudinally segregated high-elevation refugia along the

humid forests that historically spanned the island's eastern escarpment – and it reaffirms the need for continued collection and preservation of specimens in one of the world's hottest biodiversity hotspots as these forests face unprecedented rates of anthropogenic alteration.

Table of Contents

List of Figures

List of Tables

List of Supplementary Figures

List of Supplementary Tables

Acknowledgements

This dissertation has been a collaborative effort and I am sincerely grateful for the hard work and support of every individual who has been involved. First I would like to thank the members of my graduate committee – Sharon Jansa, Derek Sikes, and Naoki Takebayashi – who have been a source of encouragement throughout this project. I am especially grateful to my advisor, Link Olson, for generating most of the baseline data used in this research; for collecting morphological measurements from thousands of specimens; for making sure I had funding for my salary, conferences, and fieldwork; for being accessible when I needed him; for offering great advice about work and life; and for taking me on as his graduate student.

This research also would not have been possible without Steve Goodman, an unparalleled modern-day naturalist who collected the majority of the tenrec specimens used in this study. I am grateful to Steve for giving me the opportunity to conduct fieldwork, for being a responsive and helpful coauthor, and for sharing his wealth of knowledge about tenrecs and Madagascar.

Financial support for this dissertation was provided by a Graduate Research Fellowship from the National Science Foundation (NSF); a Graduate Student Research Award and a Publisher's Award for Excellence in Systematic Research from the Society of Systematic Biologists (SSB); a Grant-in-Aid of Research and the A. Brazier Howell Award from the American Society of Mammalogists (ASM); and a Dissertation Completion Fellowship from the University of Alaska Fairbanks (UAF) Graduate School. I was also supported by three semesters of teaching assistantships from the UAF Department of Biology and Wildlife and two semesters of research assistantships on NSF grant DEB-1120904. NSF grant DEB-1120904 also funded a large portion of the lab work and other costs for this research. Travel funding for workshops, training, and conferences was provided by SSB, the UAF Graduate School, and the UAF College of Natural Science and Mathematics.

This study would not have been possible without the generous access to and assistance with museum specimens and tissues provided by: Christopher J. Raxworthy, Nancy B. Simmons, Eileen Westwig, and Neil P. Duncan (American Museum of Natural History); Bruce D. Patterson, Lawrence R. Heaney, Adam W. Ferguson, and the late William T. Stanley (Field Museum of Natural History); Wim Van Neer (Musée Royal d'Afrique Centrale); Cécile Callou, Christiane Denys, and Violaine Colin (Muséum National d'Histoire Naturelle, Paris); Mark Omura, Judy Chupasko, and Hopi Hoekstra (Museum of Comparative Zoology, Harvard); James L. Patton, Michael Nachman, and Christopher J. Conroy (Museum of Vertebrate Zoology, University of California, Berkeley); Paula D. Jenkins, Louise Tomsett, Roberto Portela Miguez, and Richard Sabin (The Natural History Museum, London [formerly the British Museum of Natural History]); Linda K. Gordon, Darrin P. Lunde, and Michael D. Carleton

(National Museum of Natural History, Smithsonian); Daniel Rakotondravony, Steven M. Goodman, and Voahangy Soarimalala (Université d'Antananarivo, Département de Biologie Animale); and Priscilla K. Tucker, Cody W. Thompson, Steve Hinshaw, and Philip Myers (University of Michigan Museum of Zoology).

I am grateful to Voahangy Soarimalala, Daniel Rakotondravony, Achille Raselimanana, Marie Jeanne Raherilalao, and the many other staff members and students at the Vahatra Association who have done so much for conservation and education in Madagascar, and who made me feel welcome during my visits.

I would like to thank everyone in the University of Alaska Museum (UAM) Mammals Department for their support and assistance in countless areas of this project: Aren Gunderson, Kyndall Hildebrandt, Nick Kerhoulas, Jon Nations, Katie Rubin, Kendall Mills, Michelle Cason, and numerous undergraduate and high school students. I am also grateful to professors Kevin Winker, Steffi Ickert-Bond, and Devin Drown, who have helped with many of my analyses and have provided excellent advice about work and life.

I am deeply grateful to Bryan Carstens and all of the students in his Model-Based Phylogeography course at Ohio State for making me feel welcome during my visit and for continuing to give their support after I returned to Fairbanks. Thanks also to Michael Alfaro and the members of his lab at UCLA for training me in wet-bench next-generation sequencing methods, as well as Travis Glenn and Brant Faircloth for teaching me how to analyze the data. Also to Ian Herriott in the UAF Core Lab, thank you for always being willing to drop whatever you were doing to help me troubleshoot, or just to have a fun chat about DNA.

I would like to thank Bill Stanley (1956-2015), who provided support during the early development of this project. Bill made important contributions to African mammal systematics, and was one of the best science communicators I have ever known. I am grateful to Bill for inspiring hundreds of people to be excited about natural history.

To my mom, my dad, and my siblings, who are still the first people I call when I get good news at work, and who always give me their unflinching support (even though they don't always understand what I do); to Nick, Hatcher, Hurricane, and the rest of the Konefal family for making Fairbanks feel like home these last few years; and to the many friends who have supported me over the past six years: I'm afraid to list you all because I don't want to forget anyone in print, but I truly don't think I'll ever know better people. Thank you for keeping me afloat. I love you all.

And finally, to the hundreds of Malagasy scientists and nonscientists who have carried field gear, dug pitfall lines, signed permits, cooked camp food, built make-shift research stations, and executed all

other behind-the-scenes work that has been required for fieldwork over the last few decades: this dissertation would not have been possible without you. *Misaotra betsaka*.

General Introduction

Background

Madagascar has long been a source of fascination for biologists. Often called the eighth continent, Madagascar's biotic composition is quite distinct from neighboring Africa (de Wit 2003). Some taxonomic groups are completely absent (e.g., artiodactyls) while others are exceptionally diverse (e.g., chameleons). Underlying this apparent imbalance are a remarkable number of *in situ* evolutionary radiations that have taken place during Madagascar's 88 million years of isolation (Yoder and Nowak 2006). Its native, terrestrial mammals are particularly spectacular: only four extant lineages exist on the island – tenrecs (Tenrecidae), nesomyine rodents (Nesomyinae), lemurs (Lemuroidea), and carnivorans (Eupleridae) – and all are endemic, i.e., found nowhere else (Goodman and Benstead 2005; Poux et al. 2005).

Madagascar is recognized as a biodiversity hotspot partly because of its high levels of biodiversity and endemism, but also because it has experienced exceptional habitat loss (Myers et al. 2000). Since the 1950s, forested areas have decreased by almost 40%, primarily due to slash-and-burn agricultural practices and climate change (Harper et al. 2007). The rate of habitat loss is still accelerating and is currently estimated at approximately 0.93 – 2.33 % deforestation per year in the humid forests, which form a north-south belt along the eastern side of the island, and 0.46 – 1.17% per year in the dry forests on the western side of the island (Grinand et al. 2013).

Throughout these habitats live small- to medium-sized placental mammals (2-1800 g) called tenrecs (Tenrecidae) that are completely endemic to Madagascar (Everson et al. 2016). Tenrecs have been called a "remarkable adaptive radiation" (Eisenberg and Gould 1970, p. 1) or, in the words of popular science writer David Quammen (1996, p. 43), "they aren't only peculiar, they're peculiar in a profusion of different ways". There are tenrecs with exceptionally long (*Microgale longicaudata*) and short (*Tenrec ecaudatus*) tails; tenrecs highly adapted to swimming underwater (*Microgale mergulus*) or burrowing (*Oryzorictes* spp.); tenrecs that look superficially like shrews (*Microgale*), moles (*Oryzorictes*), and hedgehogs (*Echinops*, *Setifer*); and tenrecs that look like nothing else on earth (*Hemicentetes*) (Olson and Goodman 2003; Olson 2013).

For more than a century, tenrecs were grouped with shrews, moles, hedgehogs, golden moles, and solenodons in the order Insectivora and later Lipotyphla. Overwhelming molecular phylogenetic evidence (e.g., Stanhope et al. 1998; Meredith et al. 2011; O'Leary et al. 2013) has since demonstrated that tenrecs are afrotherians, meaning that they are most closely related to golden moles (Chrysochloridae), sengis (Macroscelidea), aardvarks (Tubulidentata), elephants (Proboscidea), dugongs and manatees (Sirenia), and hyraxes (Hyracoidea); thus, their superficial resemblance to shrews, moles, and hedgehogs are remarkable examples of convergence. Concurrent with changes in mammalian taxonomy at the ordinal

and superordinal levels, the alpha taxonomy of tenrecs has been rapidly changing as well; since 1992 (and prior to this dissertation), 10 new species have been described and an 11th resurrected from synonymy. This elevated rate of species discovery is largely due to three factors: increased scientific collecting in new areas of the island; increased use of pitfall traps, which have proven to be particularly effective for small-bodied tenrecs; and the application of rigorous molecular and morphometric approaches. Documenting these previously unrecognized species continues to be urgently important in light of ongoing and accelerating habitat loss on Madagascar.

Dissertation Scope

My dissertation research encompasses several fundamental areas of evolutionary biology as they pertain to Madagascar's tenrecs: (1) *phylogenetics*: the study of the evolutionary relationships (the branching patterns of shared ancestry) among individuals or lineages, (2) *taxonomy*: the practice of organizing, describing, and naming groups of organisms, (3) *phylogeography*: the study of the biotic and abiotic historical factors that have driven the recent diversification of organisms and their geographic distributions, and (4) *species delimitation*: the practice of determining the species-level taxonomic boundaries between groups of individuals, usually by analyzing genetic and/or phenotypic datasets.

In Chapter 1 (see also Everson et al. 2016), I use DNA sequence data to generate the first phylogeny of tenrecs to include every recognized species. In so doing, I resolve several long-standing taxonomic mysteries, including the phylogenetic positions of two ecologically unique species: the mouse-eared tenrec *Geogale aurita* and the web-footed tenrec *Microgale mergulus*. I also use this phylogeny as a framework for analyzing other diverse datasets. For example, by incorporating known ages from the fossil record, I am able to reconstruct the timing of colonization of Madagascar, and by incorporating geographic range data I am able to determine which habitat types are associated with increased rates of speciation.

In Chapter 2 (see also Everson et al. 2018), to better understand the finding that speciation rates are elevated in the humid forest, I conduct a phylogeographic study on a single tenrec genus (*Oryzorictes,* 2 spp.) that is endemic to, and widespread throughout, Madagascar's eastern humid-forest belt. Using a combination of genetic and morphometric data from many individuals across the genus's geographic range, I determine how many genetically and morphologically distinct groups are present and whether these groups correspond to cryptic species. Related to this research, I also test the performance of a popular method for species delimitation and address some taxonomic challenges related to a lack of holotype specimens.

In Chapter 3, I analyze mitochondrial DNA sequence data from 823 individuals representing 13 species of tenrecs plus 7 species of nesomyine rodents that are all restricted to Madagascar's eastern

humid forests. The goal of this chapter is to test the idea (proposed in Chapter 2) that high-elevation regions have played an important role in the speciation process by functioning as refugia during paleoclimatic cycles. I investigate whether population divergence is spatially and/or temporally congruent across multiple species, and whether this corresponds to four primary centers of diversity in the humid forests: the northern, central, and southern highlands, and a volcanic mountain in the far north (Montagne d'Ambre). A common phylogeographic pattern across many species suggests that the same climatic and/or geological pressures may have driven diversification in Madagascar's flora and fauna.

Finally, in Chapter 4 I continue to explore patterns and processes of diversification in Madagascar' humid forests by studying species limits and gene flow in a clade of shrew tenrecs endemic to that region that includes three currently recognized species: *Microgale fotsifotsy, M. soricoides,* and *M. nasoloi*. I ask whether or not gene flow has occurred among these species, especially the two sympatric species *M. fotsifotsy* and *M. soricoides*. I use morphometric data and a massively multi-locus genetic dataset to determine phylogenetic relationships, divergence times, and the extent of post-speciation gene flow. In the process of analyzing these morphometric and genetic datasets, I also discover that *M fotsifotsy* is, in fact, a complex of two distinct species.

Taken as a whole, the contents of this dissertation vastly improve our understanding of the tenrec branch of the tree of life, which will interest the many scientists who study island biodiversity and adaptive radiations. It also illuminates a driving force of diversification in Madagascar's humid forests and highlights the utility of integrating diverse datasets for species delimitation. These aspects of my research will be useful to the many researchers who are working to understand the origins of Madagascar's astounding and highly threatened biodiversity.

References

Eisenberg J.F., Gould E. 1970. The tenrecs: A study in mammalian behavior and evolution. Smithson. Contrib. to Zool. 27:1–137.

Everson K.M., Hildebrandt K.B.P., Goodman S.M., Olson L.E. 2018. Caught in the act: Incipient speciation across a latitudinal gradient in a semifossorial mammal from Madagascar, the mole tenrec *Oryzorictes hova* (Tenrecidae). Mol. Phylogenet. Evol. 126:74–84.

Everson K.M., Soarimalala V., Goodman S.M., Olson L.E. 2016. Multiple loci and complete taxonomic sampling resolve the phylogeny and biogeographic history of tenrecs (Mammalia: Tenrecidae) and reveal higher speciation rates in Madagascar's humid forests. Syst. Biol. 65:890–909.

Goodman S.M., Benstead J.P. 2005. Updated estimates of biotic diversity and endemism for Madagascar. Oryx 39:73–77.

Grinand C., Rakotomalala F., Gond V., Vaudry R., Bernoux M., Vieilledent G. 2013. Estimating

deforestation in tropical humid and dry forests in Madagascar from 2000 to 2010 using multi-date Landsat satellite images and the random forests classifier. Remote Sens. Environ. 139:68–80.

Harper G., Steininger M., Tucker C., Juhn D., Hawkins F. 2007. Fifty years of deforestation and forest fragmentation in Madagascar. Environ. Conserv. 34:325–333.

Meredith R., Janečka J., Gatesy J. 2011. Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. Science 334:521–524.

Myers N., Mittermeier R.A., Mittermeier C.G., da Fonseca G.A., Kent J. 2000. Biodiversity hotspots for conservation priorities. Nature 403:853–858.

O'Leary M., Bloch J., Flynn J. 2013. The placental mammal ancestor and the post–K-Pg radiation of placentals. Science 339:662–667.

Olson L.E. 2013. Tenrecs. Curr. Biol. 23:R5–R8.

Olson L.E., Goodman S.M. 2003. Phylogeny of Madagascar's tenrecs (Lipotyphla, Tenrecidae). In: Goodman S.M., Benstead J.P., editors. The Natural History of Madagascar. Chicago: University of Chicago Press. pp. 1235–1242.

Poux C., Madsen O., Marquard E., Vieites D.R., de Jong W.W., Vences M. 2005. Asynchronous colonization of Madagascar by the four endemic clades of primates, tenrecs, carnivores, and rodents as inferred from nuclear genes. Syst. Biol. 54:719–730.

Quammen D. 1996. The Song of the Dodo: Island Biogeography in an Age of Extinctions. London: Random House.

Stanhope M.J., Waddell V.G., Madsen O., de Jong W., Hedges S.B., Cleven G.C., Kao D., Springer M.S. 1998. Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. Proc. Natl. Acad. Sci. USA. 95:9967–9972.

de Wit M.J. 2003. Madagascar: Heads it's a continent, tails it's an island. Annu. Rev. Earth Planet. Sci. 31:213–248.

Yoder A.D., Nowak M.D. 2006. Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. Annu. Rev. Ecol. Evol. Syst. 37:405–431.

Chapter 1 – Multiple loci and complete taxonomic sampling resolve the phylogeny and biogeographic history of tenrecs (Mammalia: Tenrecidae) and reveal higher speciation rates in Madagascar's humid forests[1]

[1]Everson, K.M., Soarimalala, V., Goodman, S.M., & Olson, L.E. (2016). *Systematic Biology* 65: 890-909.

1.1 Abstract

The family Tenrecidae (tenrecs) is one of only four extant terrestrial mammal lineages to have colonized and diversified on Madagascar. Over the past 15 years, several studies have disagreed on relationships among major tenrec lineages, resulting in multiple reinterpretations of the number and timing of historical transoceanic dispersal events between Africa and Madagascar. We reconstructed the phylogeny of Tenrecidae using multiple loci from all recognized extant species and estimated divergence timing using six fossil calibrations within Afrotheria. All phylogenetic analyses strongly support monophyly of the Malagasy tenrecs, and our divergence timing analysis places their colonization of the island at 30-56 Ma. Our comprehensive phylogeny supports three important taxonomic revisions that reflect the evolutionary history of tenrecs: (1) we formally elevate the African otter shrews to their own family Potamogalidae, thereby rendering extant Tenrecidae entirely endemic to Madagascar, (2) we subsume the semiaquatic genus *Limnogale* within the shrew-tenrec genus *Microgale*, and (3) we re-elevate the two largest-bodied shrew tenrecs, *Microgale dobsoni* and *M. talazaci*, to the genus *Nesogale* Thomas 1918. Finally, we use recently summarized habitat data to test the hypothesis that diversification rates differ between humid and arid habitats on Madagascar, and we compare three common methods for ancestral biogeographic reconstruction. These analyses suggest higher speciation rates in humid habitats and reveal a minimum of three and more likely five independent transitions to arid habitats. Our results resolve the relationships among previously recalcitrant taxa, illuminate the timing and mechanisms of major biogeographic patterns in an extraordinary example of an island radiation, and permit the first comprehensive, phylogenetically consistent taxonomy of Madagascar's tenrecs.

1.2 Introduction

Madagascar is renowned for the strikingly high endemicity of its native biota, including 100% endemism of its 171 extant, terrestrial mammal species (Soarimalala and Goodman 2011; Mittermeier et al. 2014). About 20% of these are tenrecs (Tenrecidae), small- to medium-sized placental mammals (2-1250 g) often described as an extraordinary example of adaptive island radiation (Olson and Goodman 2003; Olson 2013). Madagascar's tenrecs are notable for a number of ecomorphological, physiological, and behavioral specializations (Fig. 1.1), including semifossoriality (*Oryzorictes*); semiaquatic carnivory (*Limnogale*); caudal prehensility (*Microgale longicaudata*); heterothermy (*Geogale*); communication via dorsal spine stridulation (*Hemicentetes* spp. and *Tenrec*); and long-term hibernation without periodic arousal (*Tenrec*) (Olson and Goodman 2003; Goodman et al. 2013; Olson 2013; Lovegrove et al. 2014). As configured by certain authors, only three tenrec species occur outside Madagascar; these are the African otter shrews [Potamogalinae; genera *Potamogale* (one species) and *Micropotamogale* (two

species)], which are semiaquatic and share some superficial morphological similarities with *Limnogale* (Benstead and Olson 2003).

Tenrecs have a tumultuous taxonomic history that has resulted in multiple reinterpretations of their evolutionary and biogeographic past, and has at times altered the historical narrative of Madagascar's notably imbalanced biotic assemblage with respect to Africa. For more than a century, tenrecs were erroneously classified alongside shrews, moles, hedgehogs, golden moles, and solenodons on the basis of a shared "primitive" morphology (Simpson 1945, p. 175). However, overwhelming molecular (e.g., Stanhope et al. 1998; Meredith et al. 2011) and, more recently, morphological and paleontological (O'Leary et al. 2013) evidence supports tenrecs as members of the superordinal clade Afrotheria, a patchwork of primarily African taxa that also includes golden moles (Chrysochloridae), elephants (Proboscidea), dugongs and manatees (Sirenia), hyraxes (Hyracoidea), sengis (Macroscelidea), and aardvarks (Tubulidentata). Tenrecs and golden moles together comprise the order Afrosoricida Stanhope et al. 1998 (a misnomer, as the clade does not include true shrews [Soricidae]).

The family Tenrecidae *sensu lato* includes at least 34 living species, all but three of which (the otter shrews) are endemic to Madagascar (Soarimalala and Goodman 2011). Three subfamilies of Malagasy tenrecs are recognized by most authorities (Bronner and Jenkins 2005). The subfamily Tenrecinae includes four genera and five species (*Tenrec ecaudatus*, *Setifer setosus*, *Echinops telfairi*, *Hemicentetes nigriceps*, and *H. semispinosus*) with spiny pelage and a larger body size than most other Malagasy tenrecids. Oryzorictinae includes two species of rice- or mole-tenrecs (*Oryzorictes hova* and *O. tetradactylus*); the semiaquatic and monotypic web-footed tenrec (*Limnogale mergulus*); and shrew tenrecs (*Microgale*). Finally, the mouse-eared tenrec (*Geogale aurita*) is currently placed in its own subfamily, Geogalinae (e.g., Bronner and Jenkins 2005, following Trouessart 1881), although it has also been previously placed in Oryzorictinae (e.g., Genest and Petter 1975).

Phylogenetic and biogeographic hypotheses for Tenrecidae are rife with debate, and no study to date has included all recognized species. Three molecular phylogenetic studies have included all currently recognized tenrec genera (Olson and Goodman 2003; Asher and Hofreiter 2006; Poux et al. 2008), and these have produced discordant topologies (Fig. 1.2). *Limnogale mergulus* has been recovered as either sister to Potamogalinae (Asher 1999) or nested within *Microgale* (Olson 1999; Olson and Goodman 2003; Poux et al. 2008). A sister relationship to Potamogalinae would necessitate more than one overwater dispersal event and would suggest that some of *Limnogale*'s aquatic adaptations are synapomorphic with those of African otter shrews (reviewed in Benstead and Olson 2003). Conversely, a nested position within *Microgale* would imply a recent, rapid evolution of a suite of aquatic adaptations and in turn render *Microgale* paraphyletic. Shrew tenrecs have undergone extensive taxonomic revision since 1992, during which time 10 new species have been described and an 11th resurrected from

synonymy. Hence, resolving the phylogenetic position of *Limnogale* and the concomitant biogeographic implications would require extensive, if not exhaustive, taxon sampling within *Microgale*.

Geogale aurita, another potential biogeographic linchpin, has also been phylogenetically recalcitrant. Three studies have recovered three different topologies with regard to its position (Fig. 1.2):

(1) Olson (1999) and Olson and Goodman (2003) used the same unpublished dataset consisting of nuclear + mitochondrial data and recovered a sister relationship between *Geogale* and the remaining Malagasy tenrecs.

(2) Asher and Hofreiter (2006) included 907 nucleotides from an exonic region of the nuclear growth hormone receptor (GHR) gene plus 126 morphological characters in a series of separate and combined analyses and recovered a nested relationship of *Geogale* within Oryzorictinae.

(3) Poux et al. (2008) analyzed 5439 nucleotides of nuclear DNA from four loci and recovered a sister relationship between *Geogale* and Oryzorictinae.

Asher and Hofreiter's (2006) analyses also recovered a close affinity between *Geogale* and the Kenyan Miocene fossils †*Erythrozootes* Butler and Hopwood 1957, †*Protenrec* Butler and Hopwood 1957, and †*Parageogale* Butler 1984, although an alternative topology placing these fossils outside the Malagasy radiation could not be rejected. Depending on *Geogale*'s position in the tenrec phylogeny, a sister relationship to these extinct African taxa may indicate more than one colonization event of Madagascar or else a back-colonization to Africa, a decidedly rare phenomenon among Malagasy terrestrial vertebrates (Yoder and Nowak 2006). Poux et al. (2008) also sequenced a *Geogale* specimen that was deeply divergent from the *Geogale* sequence of Asher and Hofreiter (2006), suggesting that *Geogale* "might contain in fact more than one species" (p. 4), an as-yet untested hypothesis that may underlie the discrepancies among published studies. That the inclusion of an unrecognized 'cryptic' species could influence higher-level phylogeny and deep biogeographic inference would be notable.

Evolutionary interpretation of tenrec diversity has been hindered by a paucity of reliable fossils. †*Parageogale*, †*Protenrec*, and †*Erythrozootes* are the only confirmed (see Seiffert et al. 2007) tenrecid fossils from the Tertiary (66-2.6 Ma); no pre-Pliocene fossil tenrecs are known from Madagascar (Asher 2010). Seiffert et al. (2007) suggested that two additional genera, †*Widanelfarasia* Seiffert and Simons 2000 and †*Jawharia* Seiffert et al. 2007, from the Eo-Oligocene of Egypt, might nest within Tenrecidae. If correct, these represent the oldest records of the family; however, the authors were unable to reject the possibility that one or both taxa represent a sister group to Afrosoricida. One additional fossil, †*Ndamathaia* Jacobs et al. 1987 (now synonymized with †*Kelba* Savage 1965), was also originally described as a tenrecid, but we follow several other authors in regarding this extinct genus as a non-tenrecid (McKenna and Bell 1997; Olson 1999; Morales et al. 2000; Asher and Hofreiter 2006; Cote et al. 2007). The impoverished pre-Quaternary Malagasy fossil record has made it difficult to produce a time-

calibrated phylogeny of Tenrecidae. Past estimates of divergence times within Tenrecidae have used only external fossil node calibrations, mostly outside Afrotheria (Douady and Douzery 2003; Poux et al. 2005; Poux et al. 2008).

Previous studies have only briefly addressed the *in situ* diversification of Malagasy tenrecs and have focused almost exclusively on the question of vicariance versus over-water dispersal in the colonization of Madagascar (e.g., Poux et al. 2005). No study to date has attempted to reconstruct ancestral habitat associations or to understand habitat-specific speciation processes. It is estimated that during the Eocene or Oligocene (the approximate period of tenrecid colonization), there was an expansion of humid forested habitats in eastern Madagascar (Samonds et al. 2013). The western region, shielded from rain-bearing weather systems coming off the Indian Ocean by the island's eastern escarpment and central high plateau, is believed to have remained relatively xeric more or less throughout the Cenozoic. Colonization time estimates, and proximity of the west coast of Madagascar to continental Africa, suggest that the first tenrec(s) likely colonized Madagascar's arid western habitats. Because the majority of Madagascar's tenrecs are restricted to the eastern humid forest today (Goodman et al. 2013), we hypothesize that the expansion of humid forests on the island led to increased rates of speciation as new niches became available.

Here we present the largest molecular phylogenetic study of Tenrecidae to date (11 genes, 9,584 nucleotides) and the first to include all recognized species. This comprehensive phylogeny provides a much-needed backbone for placing newly recognized species, which have been named at a rate of nearly one every two years since 1992. We pay particular attention to the monotypic genera *Limnogale* and *Geogale*, whose phylogenetic lability has resulted in multiple biogeographic interpretations, and we resolve the discordance in previous studies regarding unexpectedly deep molecular divergence within *Geogale*. We also estimate divergence times and evaluate the diversification process, especially with respect to habitat association. This study allows a phylogenetically consistent taxonomic revision of Tenrecidae, and improves our understanding of colonization and biogeography of Madagascar.

1.3 Materials and Methods

Data Collection

Samples newly sequenced for this study were obtained from field efforts and associated vouchered museum specimens by SMG, VS, LEO, and numerous colleagues over the past two decades (Supplementary Table 1.1). This study was conducted in strict accordance with the terms of research permits issued by authorities in Madagascar (Ministère des Forêts et de l'Environnement and Madagascar National Parks), following national laws.

We obtained fresh or frozen tissues from all nominal species of Malagasy tenrecs and three species of golden moles (*Amblysomus hottentotus*, *Chrysochloris asiatica*, and *C. stuhlmanni*). We also obtained a sample from the same *Geogale* individual sequenced by Poux et al. (2008) and thought to represent a distinct species. Despite its tentative recent recognition (Soarimalala and Goodman 2011; Goodman et al. 2013), the formal resurrection of *Microgale prolixacaudata* Grandidier 1937 from synonymy with *M. longicaudata* Thomas 1882 must await accurate delineation of species boundaries and confident placement of the holotype (Olson et al. 2004). Fresh tissues from African otter shrews were not available. Instead, dried tissue removed from older skull or skeleton specimens of these taxa were subjected to the extraction and amplification protocols detailed in Olson et al. (2005).

Individuals with fresh tissues were sequenced for mitochondrial genes NADH dehydrogenase subunit 2 (ND2), 12S ribosomal RNA (12S), and tRNA-Valine (tRNA-Val), and the following eight nuclear exons: exon 1 of alpha 2B adrenergic receptor gene (A2AB), exon 1 of aquaporin 2 gene (AQP2), exon 1 of androgen receptor gene (AR), exon 1 of brain-derived neurotrophic factor gene (BDNF), exon 11 of breast and ovarian cancer susceptibility 1 gene (BRCA1), exon 10 of growth hormone receptor gene (GHR), exon 1 of recombination activating gene (RAG1), and exon 28 of von Willebrand factor gene (vWF). We were unable to obtain clean 12S+tRNA-Val sequence data from the second *Geogale* individual (MVZ 220648) and therefore excluded it from alignments and analyses. Only the mitochondrial markers were successfully amplified and sequenced in the African otter shrews, although all nuclear sequences for one species (*Micropotamogale lamottei*) were retrieved from GenBank. The majority of the sequences in this dataset (82%) were generated for this study by the authors, while the remaining sequences were obtained from GenBank. Successful primer combinations are listed in Supplementary Table 1.2, and primer sequences are listed in Supplementary Table 1.3.

Genomic DNA was extracted from frozen or buffered tissue (spleen, muscle, or kidney) using either the animal tissue protocol in the PureGene kit (Gentra Systems, Inc.) or the tissue protocol of the QIAamp Tissue Kit (Qiagen). Amplifications were performed in 20 or 30 μL reactions containing amplification buffer (with 1.5 mM $MgCl_2$), Taq polymerase, dNTPs, ddH_2O, template DNA, and forward and reverse primers. Additional $MgCl_2$ was added in some cases for a final concentration up to 5 mM. Each PCR included a negative control to test for contamination. Thermal cycling parameters and annealing temperatures for PCR varied by gene and sample, but generally used the following conditions: 94°C for three minutes, followed by 30 cycles of 94°C for 30 seconds, 50-60°C for 30 seconds, and 72°C for 75 seconds, and a final extension of 72°C for five minutes.

Aliquots of the PCR products were electrophoresed and visualized on 1% agarose gels. In some cases, bands of the appropriate size were excised and melted in 50-500 μL ddH_2O and reamplified using the amplification or nested primers. Unquantified aliquots of PCR product (1-5 μL) were used to

sequence both forward and reverse strands. Sanger sequencing was carried out at the University of Washington High Throughput Sequencing Center, or in the Field Museum's Pritzker Laboratory for Molecular Systematics and Evolution operated with support from the Pritzker Foundation. Chromatogram outputs from both forward and reverse strands were edited by eye and assembled into contigs using Sequencher 5.1 (Gene Codes Corp., Ann Arbor, MI). Sequences were trimmed to begin at the first complete codon (for exons) or the beginning of the gene encoding 12S rRNA (5' to and contiguous with tRNA-Val, also included in analyses). Exonic sequences were manually aligned in MacClade v4.08 (Maddison and Maddison 2005) with reference to the translated amino acids (i.e., inferred insertion-deletion events [indels] were constrained to triplets). 12S and tRNA-Val sequences were manually aligned to the respective secondary structure models of Springer and Douzery (1996) and Kumazawa and Nishida (1993) and were linked in all analyses. Ambiguously aligned regions in both alignment types were excluded from analyses. To test for selection of protein-coding genes, we used a codon-based Z test of positive selection averaged across all sequence pairs using MEGA5 (Tamura et al. 2013. Analyses were conducted using the method of Nei and Gojobori (1986) and the variance of the difference was computed using 1000 bootstrap replicates.

We determined the best-fit model of nucleotide substitution for each gene using the full set of 56 possible models available in ModelTest v3.7 (Posada and Crandall 1998) under the Akaike Information Criterion (AIC). We also determined the optimal partitioning scheme for the concatenated analysis and the best-fit model of nucleotide substitution for each partition using PartitionFinder v1.1.1 (Lanfear et al. 2012). We used the Greedy algorithm in PartitionFinder to identify the optimal number of partitions from 31 possible partitions representing each codon position for the 9 coding genes plus 4 partitions representing the inferred stem and loop regions of 12S and tRNA-Val. The concatenated matrix contained 9,584 base pairs (bp). All new sequences have been deposited in GenBank (accession numbers in Supplementary Table 1.1) and alignments for each gene have been submitted to Dryad (doi:10.5061/dryad.711dc).

Phylogenetic Analyses

We performed gene-tree, concatenated, and concordance analyses under both maximum-likelihood (ML) and Bayesian Inference (BI) frameworks.

Maximum Likelihood. – Maximum-likelihood analyses were performed on each gene individually and on three concatenated datasets: all nuclear genes, all mitochondrial genes, and all genes (mitochondrial + nuclear). All ML analyses were implemented in GARLI v2.0 (Zwickl 2006) using default optimization parameters. Best-fit models of nucleotide substitution for each gene or partition were used in all analyses (methods described in previous section). Nodal support was estimated from 1000

likelihood bootstrap replicates in GARLI; each replicate was allowed to run until $-\ln L$ values converged (changing less than 0.01) for 7,500 generations. We produced a majority-rule consensus tree of the 1000 bootstrap replicates using SumTrees v3.3.1 (Sukumaran and Holder 2010) and visualized the trees in FigTree v1.3.1.

Bayesian Inference. – Individual gene and concatenated phylogenies were also estimated via Bayesian Inference in MrBayes v3.2.0 (Huelsenbeck and Ronquist 2001). Each gene or partition was assigned the best-fit model of nucleotide substitution available in MrBayes. For partition one, which included 12S and tRNA-Val, we recognized a separate data partition for stem regions using the doublet model implemented in MrBayes, which accounts for the non-independence of pairing sites. We conducted all MrBayes analyses using two independent MCMC runs (four chains each) of 10,000,000 iterations each, sampling trees every 1000 iterations. Runs were combined using LogCombiner (Drummond and Rambaut 2007) and stationarity was assessed by evaluating the likelihood scores of the MCMC chains in Tracer v1.6 (Rambaut et al. 2014) and the default MrBayes convergence diagnostics (standard deviation of split frequencies and potential scale reduction factor; Gelman and Rubin 1992). The first 20% of all trees were removed as burn-in, and the last 8,000 trees were used to construct a majority-rule consensus tree and assign Bayesian posterior probabilities (BPPs) to each node.

Reanalysis of Published Datasets with Regard to Geogale. – The relationship between *Geogale* and the remaining Malagasy tenrecs has proven difficult to resolve. Three recent studies have recovered different phylogenetic positions for *Geogale* (Fig. 1.2; Olson 2003; Asher and Hofreiter 2006; Poux et al. 2008). Poux et al. (2008) hypothesized that a possible reason for the discordance between their topology and that of Asher and Hofreiter (2006) was the presence of a cryptic species in the genus, based on highly divergent GHR sequences between the two studies. In addition to sequencing a new *Geogale* specimen for this study, we extracted DNA from the same specimen (MVZ 220648) used by Poux et al. (2008) and resequenced GHR using new primers to confirm the authenticity of the published sequence.

To test the results of previously published studies regarding the phylogenetic position of *Geogale*, we first downloaded all alignments from Poux et al. (2008)—who identified a sister relationship between *Geogale* and Oryzorictinae—and from Asher and Hofreiter (2006)—who recovered *Geogale* as nested within Oryzorictinae—and replicated their partitioned BI phylogenetic analyses to confirm repeatability. Poux et al. (2008) conducted a 9-partition MrBayes analysis (partition models listed in their Table 1.6) using two runs of 1,000,000 generations sampled every 20 iterations and a 25% burn-in. Asher and Hofreiter (2006) conducted a 3-partition MrBayes analysis (GHR nucleotide sequences, recoded GHR indels, and morphology partitions using the HKY+I+Γ, binary, and Lewis Mk models, respectively) using four runs of 1,000,000 generations sampled every 100 iterations with the first 15% of samples discarded. We also used these parameters to analyze each study's GHR alignments. We then conducted all analyses

a second time, but replaced the *Geogale* sequence with that of the other study (i.e., Poux et al.'s *Geogale* sequence data were used in Asher and Hofreiter's analyses and vice versa). Finally, all analyses were run using both *Geogale* sequences. We compared support values from all trees produced by these analyses.

Bayesian Concordance Analysis. – Topological discordance among gene trees due to incomplete lineage sorting, introgression, gene duplication, and possibly other phenomena is common (Knowles and Carstens 2007), and such processes can cause concatenation analyses to fail (Kubatko and Degnan 2007). Programs based on the multi-locus coalescent model (e.g., *BEAST, BEST) explicitly account for topological discord caused by incomplete lineage sorting, but these methods may not perform well with poor intraspecific sampling (Heled and Drummond 2010) and are therefore inappropriate for our dataset. An alternative to coalescent-based species tree programs that does not require intraspecific sampling is the Bayesian Concordance Analysis, implemented in BUCKy (Ané et al. 2007). BUCKy uses posterior tree distribution files from MrBayes analyses, and estimates both the dominant history of sampled individuals and the amount of support for each relationship (Bayesian Concordance Factors). BUCKy uses a Dirichlet process prior called the discordance factor (α), which models the degree of similarity among gene trees. No assumption is made regarding the reason for discordance.

We used the posterior distribution tree files from the MrBayes analyses of each individual nuclear gene as input files in the BUCKy analyses, and the posterior distribution tree file of the concatenated mitochondrial analysis (because these genes are linked). We conducted four BUCKy analyses using different values of α (0.1, 1, 5, and 10) to model a range of prior probabilities on the number of concordance trees, from one ($\alpha = 0.1$) to seven or more ($\alpha = 10$) concordance trees. All other parameters were set to default values, and we conducted all analyses using four concurrent MCMC chains of 1,000,000 generations with a 10% burn-in.

We also used multidimensional scaling of tree-to-tree pairwise distances to visualize the relationships among our gene-specific, concatenated, and Bayesian concordance phylogenies (Hillis et al. 2005). We used the output files from each of our MrBayes analyses to calculate pairwise Robinson-Foulds (RF) distances between all trees, using every tenth tree in the treefile after burn-in (for a total of 80 trees per partition). We then plotted the tree space (including the final tree produced from the BUCKy analysis) using a principal coordinate analysis (PCoA) of the RF pairwise distance matrix. We calculated RF distances and conducted the PCoA using the R packages phangorn and APE (Paradis et al. 2004), respectively.

Divergence Time Estimation.– We employed a relaxed molecular clock model and fossil-based node constraints to establish an evolutionary timescale and estimate divergence times within Afrotheria. Fossils were used to establish constraints on six nodes. Following Parham et al. (2012), we selected fossils using three criteria:

1) The locality and age/stratigraphic level for the fossil must be specified and published.

2) The fossil must represent the oldest known member of its respective lineage.

3) Membership of the fossil in its respective lineage must have been determined or verified by a published phylogenetic or cladistic analysis.

We used fossils to place minimum age constraints on six nodes representing major splits between Afrotheria lineages (Table 1.1; labeled A-F on Fig. 1.4). Each fossil represents the oldest recognized member of its respective order or family and was therefore used as a minimum age constraint for the origin of that clade. In order to use these calibration points, the complete DNA dataset was supplemented with GenBank sequences from single representatives of the remaining afrotherian orders except for Macroscelidea, for which representatives of the divergent subfamilies Rhynchocyoninae and Macroscelidinae (giant elephant shrews and soft-furred elephant shrews, respectively; Corbet and Hanks 1968) were included (Supplementary Table 1.4).

We conducted the fossil-calibrated divergence time analysis using BEAST v1.7.5 (Drummond and Rambaut 2007). Partitions were established as in previous analyses, and we allowed independent evolution of branch lengths under the uncorrelated relaxed lognormal clock model and the Yule prior. We employed two independent MCMC runs (four chains each) of 10,000,000 iterations each, sampling trees every 1000 iterations to yield a total of 10,000 trees. Runs were combined in LogCombiner, the first 20% of these trees were removed as burn-in, and the last 8,000 trees were used to construct a majority-rule consensus tree and assign BPPs to each branch. Results were summarized with TreeAnnotator v1.6.1 (Drummond and Rambaut 2007).

We used the resulting ultrametric tree to generate a lineages-through-time (LTT) plot using the command LTT.plot in the R package APE. We then fit our data to the birth-death model and to the Yule (birth-only) model using the commands make.bd and constrain in the R package diversitree (FitzJohn 2012), and we tested the fit of the birth-death and Yule models using a likelihood ratio test. Finally, we simulated 1000 trees under the best-fit model and plotted these on our LTT plot.

Biogeographic Analysis

To test the hypothesis that Madagascar's humid-forest habitat is associated with higher rates of speciation in tenrecs, we used the geographic-state speciation and extinction (GeoSSE) model (Goldberg et al. 2011) as implemented in the R package diversitree. The GeoSSE model is an extension of the binary-state speciation and extinction (BiSSE) model (Maddison et al. 2007), which tests whether speciation and extinction rates vary as a function of a binary character. Unlike BiSSE, which accepts only binary characters, GeoSSE allows each species to occur in one or more designated habitats or areas. In this study, humid (H), dry (D), and a combination (HD) of habitat associations were specified for each

species based on comprehensive occurrence data and habitat types presented in Goodman et al. (2013; Supplementary Table 1.5). Humid and dry habitats form nearly contiguous regions along the eastern and western portions of Madagascar, respectively. The resulting GeoSSE model included seven parameters: speciation rate in states H, D, and HD (sH, sD, and sHD, respectively); extinction rate in states H and D (xH and xD); and dispersal from H to D or D to H (dH and dD).

Within GeoSSE, we tested a set of 10 models using the ultrametric tree from our BEAST analysis with otter shrews and other non-tenrecs excluded. We first tested the full model, where all seven parameters were free to vary. We then tested eight constrained subsets of the full model and compared models using both the likelihood ratio test and AIC:

1. sH, sD, sHD, xH, xD, dH, dD (full model, 7 parameters)
2. sHD=0 (no intermediate speciation, 6 parameters)
3. xH=0, xD=0 (no extinction, 5 parameters)
4. sH=sD (speciation equal between regions, 6 parameters)
5. xH=xD (extinction equal between regions, 6 parameters)
6. dH=dD (dispersal equal between regions, 6 parameters)
7. xH=0, xD=0, sH=sD (no extinction, equal speciation between regions, 4 parameters)
8. xH=0, xD=0, dH=dD (no extinction, equal dispersal between regions, 4 parameters)
9. xH=0, xD=0, sH=sD, dH=dD (no extinction, equal speciation between regions, equal dispersal between regions, 3 parameters)

In GeoSSE, we first conducted ML parameter estimation and model comparison, then used the top-ranked model to obtain more accurate parameter estimates via BI. For the BI analysis, we used ML parameter estimates as starting values and re-estimated parameter values via an MCMC analyses with 10,000 generations with the first 10% removed for burn-in. Posterior probability distributions for the GeoSSE parameters were visualized using the function profiles.plot in diversitree. We tested whether the rate of speciation was greater in humid regions than dry regions by computing the proportion of humid speciation rate values that were greater than dry speciation rate values, following Goldberg et al. (2011).

Another goal of our biogeographic analyses was to infer the habitat associations of ancestral tenrecs. To accomplish this we first performed an ancestral character state reconstruction (ACSR) of habitat types using the function asr.marginal in the package diversitree, which calculates ancestral character states using the parameter values estimated in GeoSSE. Because the GeoSSE model may not include all of the parameters relevant in our system (see Discussion), we performed two additional ACSR analyses for comparison: a BioGeoBEARs analysis, which uses the Dispersal-Extinction-Cladogenesis (DEC) model of range evolution (Ree et al. 2005; Matzke 2013), and a traditional ACSR, which uses a

standard ML approach without explicit biogeographic parameters. Both analyses used the ultrametric BEAST phylogeny and the same habitat trait matrix used in the GeoSSE analysis.

The DEC model is a geographic range evolution model in which range expansion occurs by dispersal events, range contraction occurs by local extinction events, and the probability of each event is proportional to branch length. Unlike GeoSSE, the DEC model assumes that speciation and extinction are independent of the range evolution process. The R package BioGeoBEARS provides a flexible framework for comparing biogeographic models, including the DEC model and the DEC+j model, which incorporates founder-event speciation by allowing dispersal without range expansion (the "jump" parameter, j; Matzke 2014). We reconstructed ancestral areas under both the DEC and DEC+j models and compared the fit of the two models with respect to the addition of the "j" parameter by evaluating delta-AIC and a standard likelihood-ratio test.

Finally, we performed a traditional (non-biogeographic) ML ACSR on habitat types using the Ancestral Character Evolution (ACE) function in the R package APE. In this analysis, the character state "both habitats" (HD) is a discrete third state, and there are no explicit biogeographic parameters. We tested three alternative transition-rate models: Equal Rates (ER; assumes equal transition rates among all habitats), All Rates Different (ARD; assumes different transition rates among all habitats in all directions), and Symmetrical Rates (SYM; assumes that transition rates to and from the same habitats are equal). We selected the model with the best fit to the data using a standard likelihood-ratio test, and then used the top-ranked model to reconstruct ancestral nodes. Finally, we repeated all ACSR analyses (GeoSSE, BioGeoBEARs, and ML) with African otter shrews included, because the inclusion of outgroups may influence the node reconstructions, especially at the earliest nodes (all otter shrews are endemic to humid habitats on continental Africa and were thus coded as "humid").

1.4 Results

Sequence Characteristics

The complete concatenated analysis includes four data partitions, identified by PartitionFinder (Table 1.2). Best-fit models of nucleotide substitution for each partition were identified by PartitionFinder, while best-fit models for individual genes were determined using ModelTest (Supplementary Table 1.6). All P-values from our Z-tests for selection were non-significant ($P>0.05$), indicating that none of our coding genes are under positive selection.

Phylogenetic Analyses

We generated a majority-rule consensus tree from our concatenated analysis of the 11-gene (4-partition) dataset (Fig. 1.3). Monophyly of the extant Malagasy tenrecs is recovered in all gene trees

(Supplementary Figs. 1.1 and 1.2) and all combined analyses. The concatenated analyses also recover monophyly of all tenrecid subfamilies and genera, with the exception of *Microgale*; the aquatic tenrec *Limnogale* is nested within *Microgale* in all gene trees and concatenated analyses with high support. Gene-tree topologies disagree with regard to *Geogale*'s position (Supplementary Figs. 1.1 and 1.2). It is variably recovered as sister to *Microgale* (12S + tRNA-Val), sister to Tenrecinae (ND2), sister to *Oryzorictes* (RAG1), or in a polytomy with two or more oryzorictine taxa (AQP2, BDNF, and GHR). Four out of 10 gene trees (A2AB, AR, BRCA1, and vWF) support a sister relationship between *Geogale* and Oryzorictinae, and the same relationship is recovered in the full concatenated analysis with high support.

We identified several indels in six of our genes, including a number of important molecular synapomorphies for previously contested clades (Supplementary Fig. 1.3). These include two single-codon indels found in *Microgale* and *Limnogale* (A2AB and BRCA1), a 4-codon deletion in *M. dobsoni* and *M. talazaci* (BRCA1), and a 9-codon deletion in *Microgale* and *Limnogale* to the exclusion of *M. dobsoni* and *M. talazaci* (BRCA1).

Reanalysis of Published Datasets with Regard to Geogale

We reanalyzed the datasets published by Poux et al. (2008) and Asher and Hofreiter (2006) and recovered their respective topologies. However, the position of *Geogale* varies according to which *Geogale* GHR sequence is used in the analysis (Supplementary Fig. 1.4). Specifically, a sister relationship between *Geogale* and Oryzorictinae is always recovered when Poux et al.'s (2008) *Geogale* GHR sequence (GenBank accession: AM905347) is used in the analysis, while a sister relationship between *Geogale* and *Microgale* is always recovered when Asher and Hofreiter's (2006) GHR sequence (DQ202287) is employed. When both GHR sequences are used, all analyses recover a sister relationship between *Geogale* and Oryzorictinae. Those relationships hold true for both GHR-only and multi-partition analyses.

Bayesian Concordance Analysis

The primary concordance tree produced by BUCKy (Fig. 1.3) was topologically congruent with our concatenated phylogeny. Varying the concordance prior α had no effect on topology or concordance factor (CF) values. Many splits have low CF values (see Discussion); nonetheless, our PCoA of pairwise RF distances (Supplementary Fig. 1.5) shows a close proximity between the BUCKy concordance tree and our concatenated trees, and among all fast-evolving loci (mitochondrial genes, BRCA1, GHR, RAG1, and vWF). The PCoA plot also shows a clear separation of the gene AQP2 from all other loci. This is not particularly surprising, as AQP2 contained few informative sites (Supplementary Table 1.6) and almost

no well-resolved nodes within Malagasy tenrecs (Supplementary Figs. 1.1 and 1.2). Similarly, this plot shows that the phylogenies of the slowly evolving nuclear genes BDNF, A2AB, and AR differ from the final concatenated and concordance trees.

Divergence Time Analysis

We generated a chronogram using the last 10,000 trees from our fossil-calibrated BEAST analysis (Fig. 1.4). We summarized divergence dates for all major clades (Table 1.3), and plotted prior and posterior age distributions of each fossil-calibrated node (Supplementary Fig. 1.6). The origin of Tenrecidae *sensu lato* [i.e., including African otter shrews (Potamogalinae)] was estimated in the Paleocene to early Eocene [58.92 Ma (95% CI: 66.19-50.83)], with African otter shrews diverging from Malagasy tenrecs during the Eocene [47.51 Ma (95% CI: 55.61-40.73)]. The crown diversification of Malagasy tenrecs began during the late Eocene to early Oligocene [34.94 Ma (95% CI: 41.75-29.57)]. The clade composed of *Microgale dobsoni* and *M. talazaci* diverged from the lineage leading to the remaining shrew tenrecs (plus *Limnogale*) during the Miocene [19.35 Ma (95% CI: 15.84-23.12)].

Our estimate for the divergence of African otter shrews and Malagasy tenrecs matches that of Poux et al. (2008; 55-40 Ma), although we obtained different divergence time estimates for other splits. For example, our results show a younger time for the divergence of Tenrecidae and Chrysochloridae from their MRCA (59 Ma compared to 69 Ma) and generally older times for splits within the Malagasy tenrecs (e.g. 30 Ma compared to 24 Ma for the split between Geogalinae and Oryzorictinae). Nonetheless, all of our 95% credibility interval (CI) values overlapped with those of Poux et al. (2008).

We used our BEAST tree to test two diversification models, the birth-death model and the Yule (birth-only) model, and found that the inclusion of an extinction parameter in the birth-death model did not result in a significant improvement over the birth-only (Yule) model [LnL_{Yule}=-30.23, LnL_{BD}=-30.44, not significant (p-value=0.90)]. We generated an empirical LTT plot, and plotted lineages through time for 1000 trees simulated under the Yule model (Supplementary Fig. 1.7).

Biogeographic Analysis

The top-ranking GeoSSE model had zero extinction and equal rates of dispersal between both humid and dry regions, but different rates of speciation among all habitat types, for a total of four parameters: sH, sD, sHD, and d (Table 1.4). We used this GeoSSE model in a BI analysis and plotted the posterior densities of dispersal rate and habitat-specific speciation rates (Fig. 1.5). The rate of speciation in humid habitats is nearly three times higher than in dry habitats (Fig. 1.5, Table 1.4, ANOVA p-value < 0.001) and is highest for eurytopic species (sHD = 0.217), although the posterior density curve is very diffuse (Fig. 1.5b). Finally, our ACSR analysis recovered a eurytopic MRCA of Malagasy tenrecs when

African otter shrews were excluded (85.9% both habitats, 7.7% humid habitat, 6.4% dry habitat; Fig. 1.6a), but a dry MRCA of Malagasy tenrecs when African otter shrews were included (59.1% dry habitat, 11.6% humid habitat, 29.3% both habitats; Supplementary Fig. 1.9a).

In our BioGeoBEARS analyses, the addition of the "j" parameter for founder-event speciation significantly increased the likelihood of the DEC model (DEC lnL = -26.51, DEC+j lnL = -20.44, $P < 0.001$; delta-AIC=10.14). This analysis recovered a eurytopic MRCA of Malagasy tenrecs when African otter shrews were excluded (59.2% both habitats, 28.9% humid habitat, 11.9% dry habitat; Fig. 1.6b), but a humid MRCA when African otter shrews were included (52.6% humid habitats, 29.4% dry habitats, 18.0% both habitats; Supplementary Fig. 1.9b).

Finally, in our traditional (non-biogeographic) ML analysis, neither the three-rate SYM model nor the six-rate ARD model resulted in a significant likelihood increase over the one-rate ER model [ER lnL = -22.9, SYM lnL = -21.2 (p-value=0.17), ARD lnL = -20.5 (p-value=0.42)]. The MRCA of Malagasy tenrecs was recovered in the humid habitat when only Malagasy tenrecs were considered (70.0% humid habitat, 15.5% dry habitat, 14.4% eurytopic; Supplementary Fig. 1.10a) and when African otter shrews were included (73.1% humid habitat, 13.6% dry habitat, 13.3% eurytopic; Supplementary Fig. 1.10b).

1.5 Discussion

Past studies of Tenrecidae have disagreed on the monophyly of, and relationships within, Malagasy tenrecs; these differences have significant biogeographic and evolutionary implications. Much of the conflict may be attributable to insufficient taxon and/or character sampling or to topological discordance between molecular and morphological data. Our study represents the largest genetic dataset to date to address the evolutionary history of tenrecs, and is the first study to use comprehensive taxon sampling. We mitigated the impacts of incomplete lineage sorting and conflicting gene-tree topologies by including multiple genes from both nuclear and mitochondrial genomes in a Bayesian Concordance Analysis.

Phylogenetic position of Geogale

We reanalyzed the datasets of Asher and Hofreiter (2006) and Poux et al. (2008) and found that the phylogenetic position of *Geogale* varies according to which sequence is used. When both specimens are included, a sister relationship between *Geogale* and Oryzorictinae is recovered by both Asher and Hofreiter's (2006) and Poux et al.'s (2008) analyses, which is the same relationship recovered by our concatenated and concordance analyses. All combined analyses, regardless of topology, recover a short

branch length between the *Geogale*, *Microgale* (plus *Limnogale*), and *Oryzorictes* crown diversification nodes, indicating that these three lineages likely diverged rapidly.

Our results are consistent with the suggestion by Poux et al. (2008) of two cryptic species of *Geogale*. Genetic distances between the two *Geogale* specimens are on par with those observed between other recognized species (Supplementary Fig. 1.8), yet the two individuals always form a sister pair. Future work should investigate both geographical and morphological differences between the two purported taxa. It would not be surprising if an undescribed species of *Geogale* exists, particularly in light of the scant taxonomic attention the genus has received, the few specimens available until recent years, and the recent descriptions of previously unknown, similarly sized *Microgale* species.

Nonmonophyly of Microgale

This study supports the contention of Guth et al. (1959) that *Limnogale* is "simply an aquatic *Microgale*" (p. 447; translated from the original French). All searches, including all individual gene analyses, nested *Limnogale* within the shrew tenrecs with high support, rendering *Microgale* paraphyletic. We therefore follow Olson and Goodman's (2003) recommendation that *Limnogale* be considered a junior synonym of *Microgale* (Table 1.5; see also Asher and Helgen 2010) but we continue to use the generic name *Limnogale* throughout the remainder of this discussion for the sake of clarity and continuity.

The nested position of *Limnogale* within *Microgale* contradicts Asher's (1999) results favoring a sister relationship to the African otter-shrews and Eisenberg's (1981) speculation of a sister relationship to the remaining Malagasy tenrecs. Asher (1999) identified two cranial synapomorphies to support a Potamogalinae + *Limnogale* clade – a fenestrate basioccipital and absence of a lacrimal foramen – but Olson and Goodman (2003) considered both characters dubious, highlighting the presence of a fenestrate basioccipital in other shrew tenrecs not included in Asher's (1999) matrix and contradicting the claimed absence of a lacrimal foramen in *Limnogale* (but see Asher and Hofreiter 2006).

The position of *Limnogale* within *Microgale* is less certain. Most of our slowly evolving nuclear genes were unable to resolve relationships within *Microgale*, which is the most recent and rapid tenrec radiation; these genes recover a polytomy between *Limnogale* and four or more *Microgale* species. However, four genes and all multi-locus analyses support a sister relationship with *M. parvula*. This result is surprising, as *L. mergulus* and *M. parvula* represent the largest (60-107 g) and smallest (adult mass 2.0-5.0 g) members of the clade, respectively (Soarimalala and Goodman 2011; Olson unpubl.). *Limnogale*'s large body size compared to the shrew tenrecs likely reflects the general trend for aquatic small mammals to be larger than their terrestrial sister species (reviewed in Benstead and Olson 2003).

Revised and phylogenetically consistent taxonomy of extant tenrecs

Our inclusive taxon sample and well-resolved phylogeny allow for a revised taxonomy of extant tenrecs reflective of their evolutionary history (Table 1.5). Perhaps most notably, we formally re-elevate the African otter shrews (formerly Potamogalinae) to familial rank (Potamogalidae). Allman's (1865, 1866) decision to name Potamogalidae in his description of *Potamogale velox* stemmed not from its distinctiveness from tenrecids but its perceived similarities to *Solenodon*, a Caribbean genus long allied with tenrecs and African otter shrews (Dobson 1882; Simpson 1945; McDowell 1958) but now placed in a separate superordinal clade based analyses of molecular, morphological, and paleontological data (Roca et al. 2004; O'Leary et al. 2013). Potamogalidae has been recognized as the sister family to Tenrecidae *sensu stricto* by a number of authors (e.g., Simpson 1945; Eisenberg 1981; Salton and Szalay 2004; Seiffert et al. 2007), and no explicit or compelling opposition to this arrangement has been expressed in the recent literature. Morphologically, African otter shrews are distinct from all Malagasy tenrecs, and every other taxon with which they have been allied, in lacking clavicles and possessing syndactylous (conjoined) second and third hindfoot digits (reviewed in Guth et al. 1959; Guth 1960; Olson 1999). Double-rooted canines and absent or indistinct hindfoot thenara (integumental pads) also set African otter shrews apart from tenrecs (Olson 1999), as does the aforementioned absence of a lacrimal foramen (Olson and Goodman 2003, but see Asher and Hofrieter 2006). Collectively, these differences, in addition to the inferred antiquity of the potamogalid lineage (Meredith et al. 2011; this study), seem more than sufficient to warrant its recognition as a family distinct from Malagasy tenrecs.

Within Tenrecidae, our results support the three traditional subfamilies recognized by most recent authors (e.g., Bronner and Jenkins 2005; Soarimalala and Goodman 2011; Goodman et al. 2013). *Geogale* has at times been placed in Oryzorictinae (Cabrera 1925; Simpson 1945; Van Valen 1967), but most recent classifications have recognized Geogalinae (McKenna and Bell 1997; Bronner and Jenkins 2005). This is supported by its long history of evolutionary independence (~30 Myr) and a suite of traits that render *Geogale* distinct from Oryzorictinae, including unique tarsal morphology (Salton and Szalay 2004), a reduced dental formula, a specialized diet consisting primarily of termites, and heterothermy across a wide range of ambient temperature (reviewed in Stephenson 2003).

In addition to the synonymization of *Limnogale* with *Microgale*, we propose another revision to shrew tenrec taxonomy. The genus *Nesogale* was erected by Thomas (1918) to accommodate *M. dobsoni* and *M. talazaci*, both of, which are much larger and more robust than other shrew tenrecs. He also reported a "peculiar sinuosity" in the skull profile, and reduced or absent secondary cusps on the incisors and canines. MacPhee (1987) noted that *M. dobsoni* and *M. talazaci* are the only shrew tenrecs without premolar diastemata; he nonetheless considered *Nesogale* a junior synonym of *Microgale*.

Our results support the resurrection of *Nesogale*, as all of our concatenated and species-tree analyses recover reciprocal monophyly between this grouping and *Microgale* (including *Limnogale*). Nine out of 10 individual gene trees also support monophyly of *Nesogale*. We also identified a 4-codon deletion in BRCA1 (beginning at alignment position 826) unique to *Nesogale*, and a 9-codon deletion in BRCA1 (alignment position 889) unique to *Microgale* exclusive of *Nesogale*. Morphologically, both species are readily diagnosable from all but one *Microgale* on the basis of their relatively enlarged, canine-like ('caniniform') lower 2nd incisors (Olson 1999; Jenkins 2003), the aforementioned absence of premolar diastemata (MacPhee 1987), and generally large body size (Soarimalala and Goodman 2011). While they share all these features with *Limnogale mergulus*, the latter is easily distinguished from all other Malagasy tenrecs by its numerous semiaquatic adaptations (Benstead and Olson 2003).

Some authors have argued that Tenrecomorpha has priority over Afrosoricida as the ordinal name for tenrecs, otter shrews, and golden moles (e.g., Mouchaty et al. 2000; Malia et al. 2002; Asher 2010). We agree with the rationale presented by Bronner and Jenkins (2005) (while sharing their misgivings) for recognizing Afrosoricida.

Bayesian Concordance Analysis

Our Bayesian Concordance Analysis (BUCKy) recovered the same topology as our concatenated analysis (Fig. 1.3); however, a number of strongly supported clades in the concatenated tree have low CF values in the primary concordance tree. Although CF values are superficially similar to bootstrap percentages or Bayesian posterior probabilities, they should not be interpreted as measures of statistical uncertainty (Baum 2007); rather, they represent the proportion of gene regions that contain a particular split (Ané et al. 2007). Low CFs may be caused by low levels of phylogenetic information in particular genes, as was the case in AQP2, BDNF, and A2AB here (Weisrock 2012). Where CF values on our primary concordance tree are especially low (<0.50), it may still be inappropriate to interpret these values as poor support for that clade. As an example, consider the Geogalinae-Oryzorictinae split, which received a CF value of 0.499. We can say that this is the best-supported relationship for Geogalinae, even though it is reflected in less than half of our dataset, because all alternative branching scenarios for Geogalinae were recovered with much lower CF values: a sister relationship between Geogalinae and *Microgale* received a CF of 0.148 (0.000,0.222), a sister relationship between Geogalinae and *Oryzorictes* received a CF of 0.142 (0.111,0.222), and a sister relationship between Geogalinae and Tenrecinae received a CF of 0.125 (0.111,0.222).

Divergence Timing

We used six afrotherian fossil-based node calibration points in our divergence time analysis. By contrast, Douady and Douzery (2003), Poux et al. (2005), and Poux et al. (2008) each used only one node calibration point within Afrotheria [†*Phosphatherium* (Gheerbrant et al. 1996), used to calibrate the Paenungulata node], plus several non-afrotherian fossils. Nonetheless, our divergence time estimates are generally on par with those of previous studies (Table 1.3; Table 1 in Douady and Douzery 2003; Table 2 in Poux et al. 2005; Table 3 in Poux et al. 2008). The specific timing of Malagasy colonization is still uncertain because dispersal may have occurred at any point along the branch leading to crown Malagasy tenrecs. We therefore conservatively conclude that tenrecs arrived on Madagascar between 55.6 and 29.6 Ma, where 55.6 Ma is the oldest value in the 95% CI for the divergence of African otter shrews and Malagasy tenrecs and 29.6 Ma is the youngest possible time in the 95% CI for the crown diversification of Malagasy tenrecs. These dates overlap broadly with the colonization timing estimates reported by Douady et al. (2002; 55-37 Ma), Poux et al. (2005; 50.3-19.7 Ma) and Poux et al. (2008; 51-26 Ma).

Biogeographic Analysis

We found higher rates of diversification in humid versus xeric habitats (Fig. 1.5). Additionally, our top-ranking GeoSSE models do not parameterize extinction (Table 1.4), suggesting that extinction may have been relatively rare over the evolutionary history of tenrecs. This corresponds well with our diversification model results, which showed that a Yule (birth-only) model best explains the observed diversification pattern in tenrecs (Supplementary Fig. 1.7). A recent study found that BiSSE has a high type I error rate, meaning that neutral traits may falsely appear to be positively correlated with speciation, and we can assume that other – SSE methods (including GeoSSE) may also be susceptible to this problem (Rabosky and Goldberg 2015); however, the increased rate of humid habitat speciation recovered in this study does not seem unreasonable. This result matches worldwide patterns of increased species diversity in humid tropical forests (Gaston 2000).

Our reconstruction of ancestral habitats is the first such attempt for any of Madagascar's mammal radiations. Over two-thirds of extant Malagasy tenrec species (n=23) are currently restricted to humid environments. Another two species occupy both humid and arid habitats (Supplementary Table 1.5). This unevenness, which is not associated with habitat-biased small mammal inventories over the past few decades (Goodman et al. 2013, Fig. 1, p. 212), is most pronounced in the subfamily Oryzorictinae (shrew, mole, and aquatic tenrecs), of which over three-quarters (21 out of 25) are known only from humid areas. Unsurprisingly, the MRCA of Oryzorictinae was recovered as a humid-forest endemic. Ancestral habitats within Tenrecinae (spiny tenrecs), on the other hand, were not confidently reconstructed. Of the five

tenrecine species, two occupy humid habitats, one occupies dry habitats, and two are eurytopic; therefore, estimates at the MRCA of Tenrecinae from all ACSR models were ambiguous.

The ancestral habitat results varied substantially depending on which reconstruction model was used. For example, a defining feature of the GeoSSE model is that transition from one habitat to another must occur via an intermediate "widespread" step; thus, in all cases where sister taxa occupy different habitats, their ancestor is recovered as widespread with high probability (Fig. 1.6a). This is not an assumption of the other two models, and "widespread" was therefore not always recovered. Results also varied depending on whether African otter shrews were included. All three otter shrew species occupy humid habitats on continental Africa, and their inclusion generally increased support for a humid-forest MRCA of tenrecs plus otter shrews as well as the MRCA of Malagasy tenrecs (Supplementary Figs. 1.9b and 1.10).

The most notable examples of disparity among the three models involve the deepest nodes on the phylogeny. Both GeoSSE and BioGeoBEARs reconstructed a widespread MRCA of Malagasy tenrecs, while the standard ML analysis reconstructed a humid-forest ancestor. Results from the standard ML analysis are perhaps the least reliable, as this model is not explicitly biogeographic – it does not contain any biogeographic parameters and "both habitats" is a distinct third state rather than a combination of humid and dry. Conversely, both the GeoSSE model and BioGeoBEARs' DEC+j model contain parameters that are important to consider in this system. GeoSSE allows non-independence of speciation rate with respect to habitat type, while DEC+j allows jump dispersal. In GeoSSE, jump dispersal will manifest as spuriously high values of sHD, which may have occurred in this study (Fig. 1.5; Goldberg et al. 2011). Furthermore, BioGeoBEARs offers a distinct advantage in ACSR by reconstructing ancestral areas as a triplet of parent and daughter states at each node. Results from GeoSSE, where only the nodal states are reconstructed, may be confusing by comparison. An improved biogeographic model would contain elements from both the GeoSSE and BioGeoBEARS models.

Ultimately, there is a two-fold problem with any method for reconstructing ancestral habitat association: we do not know if present-day habitats existed in the past (or if they existed in the same geographical area), nor do we know if species exhibit niche conservatism or if they have adapted to changing habitats over time (Donoghue and Edwards 2014). Our analyses variably recovered an ancestral Malagasy tenrec that occupied humid, dry, or both habitats, but we do not know if any of these habitat types existed at that time. Wells (2003) and Samonds et al. (2013) agree that humid forests likely did not arise on Madagascar until the Oligocene or early Miocene (after colonization by tenrecs), coincident with the commencement of Madagascar's current monsoon pattern. Before this time, it is likely that Madagascar was much drier, or that it was covered by habitats that do not exist today. The ancestral state reconstructions at these deeper nodes on our tree should therefore be interpreted with caution.

A commonality among all ACSR analyses of our data is that they require at least three independent transitions from humid to dry habitats in *Microgale*. Notably, three of the four arid-adapted shrew tenrecs were only discovered and described in the past two decades. This underscores the need for continued taxonomic and inventory work, particularly in undersampled regions of Madagascar, not only to document more thoroughly its biodiversity but also to better understand the underlying processes that generated it. The recent (Olson et al. 2009) description of *M. grandidieri*, for example, and its sister relationship to *M. brevicaudata* suggests the potential for post-colonization diversification by shrew tenrecs in arid habitats as opposed to a 'dead end' model of peripheral isolation in suboptimal habitat.

Our study resolves relationships among extant tenrecs, provides a chronology of their diversification on Madagascar, and suggests an important historic role of the island's now greatly reduced and threatened humid forests (Harper et al. 2007) as a potential species pump. It also highlights the importance of comprehensive taxon sampling—even among closely related species—in higher-level phylogenetic and biogeographic analyses. Future studies employing intraspecific sampling and niche modeling would help identify factors driving continued *in situ* diversification in tenrecs and might possibly shed further light on historic processes. Further, fine-level phylogeographic studies are needed to investigate population differentiation within widely distributed taxa, which in turn should provide insight into patterns of speciation and elucidate undetected cryptic species. These additional studies should incorporate Madagascar's changing climate and landscape, not only to better understand the biogeographic history of tenrecs, but also to predict the impacts of future change.

1.6 Acknowledgements

We dedicate this study to our beloved friend and colleague William T. Stanley (1956-2015), whose support and encouragement at every stage of this project improved it immeasurably. His untimely death represents a profound loss to the field of mammalian systematics.

We wish to acknowledge Madagascar National Parks and the Ministère des Forêts et de l'Environnement for providing authorizations for the capture, collection, and exportation of specimens under a protocol of collaboration between the Département de Biologie Animale of the Université d'Antananarivo, Association Vahatra, and the Field Museum of Natural History.

We thank reviewers Frank E. Anderson, Erika J. Edwards, Felipe Zapata, and an anonymous reviewer for their thorough and constructive comments. We are deeply grateful to the following curators and collection managers and their respective institutions for access to tissues and specimens: Christopher J. Raxworthy, Nancy B. Simmons, and Darrin P. Lunde (American Museum of Natural History); Paula D. Jenkins and Richard Sabin (The Natural History Museum, London [formerly the British Natural History Museum); Bruce D. Patterson, Lawrence R. Heaney, and William T. Stanley (Field Museum of Natural

History); Wim Van Neer (Musée Royal d'Afrique Centrale); James L. Patton, Michael Nachman, and Christopher J. Conroy (Museum of Vertebrate Zoology, University of California, Berkeley); Daniel Rakotondravony (Université d'Antananarivo, Département de Biologie Animale); and Michael D. Carleton and Linda Gordon (National Museum of Natural History, Smithsonian). We thank Valérie Reeb for translating Guth et al. (1959, 1960) into English and Kyndall B.P. Hildebrandt for generating much of the DNA sequence data. Finally, we express our deepest gratitude to the following colleagues for their feedback and support over the epic course of this project: Michael D. Alfaro, Robert J. Asher, J. William O. Ballard, F. Keith Barker, John Bates, Michelle M. Cason, Barry Chernoff, Amy C. Driskell, John J. Flynn, Shannon J. Hackett, Lawrence R. Heaney, Sharon A. Jansa, Bruce D. Patterson, Daniel Rakotondravony, Eric J. Sargis, Erik R. Seiffert, Lee Weigt, Neal Woodman, and Anne D. Yoder.

1.7 References

Allman G.J. 1865. [no title]. Proc. Zool. Soc. Lond. 1865:467.

Allman G.J. 1866. On the characters and affinities of *Potamogale*, a new genus of insectivorous mammals. Trans. Zool. Soc. London 6:1–16.

Ané C., Larget B., Baum D.A., Smith S.D., Rokas A. 2007. Bayesian estimation of concordance among gene trees. Mol. Biol. Evol. 24:412–426.

Asher R.J. 1999. A morphological basis for assessing the phylogeny of the "Tenrecoidea" (Mammalia, Lipotyphla). Cladistics 15:231–252.

Asher R.J. 2010. Tenrecoidea. In: Werdelin L., Sanders W.J., editors. Cenozoic mammals of Africa. Berkeley: University of California Press. p. 99–105.

Asher R.J, Helgen K.M. 2010. Nomenclature and placental mammal phylogeny. BMC Evol. Biol. 10:102.

Asher R.J., Hofreiter M. 2006. Tenrec phylogeny and the noninvasive extraction of nuclear DNA. Syst. Biol. 55:181–194.

Baum D.A. 2007. Concordance trees, concordance factors, and the exploration of reticulate genealogy. Taxon 56:417–426.

Benoit J., Orliac M., Tabuce R. 2013. The petrosal of the earliest elephant-shrew *Chambius* (Macroscelidea: Afrotheria) from the Eocene of Djebel Chambi (Tunisia) and the evolution of middle and inner ear of elephant-shrews. J. Syst. Palaeontol. 11:907–923.

Benstead J.P., Olson L.E. 2003. *Limnogale mergulus*. In: Goodman S.M., Benstead J.P., editors. The natural history of Madagascar. Chicago: University of Chicago Press. p. 1267–1273.

Bronner G., Jenkins P. 2005. Afrosoricida. In: Wilson D.E., Reeder D.M., editors. Mammal species of the world: a taxonomic and geographic reference. Baltimore: Johns Hopkins University Press. p. 71–81.

Butler P.M. 1972. The problem of insectivore classification. In: Joysey K., Kemp T., editors. Studies in vertebrate evolution. Edinburgh: Oliver and Boyd. p. 253–265.

Butler P.M. 1984. Macroscelidea, Insectivora, and Chiroptera from the Miocene of East Africa. Palaeovertebrata 14:117–200.

Butler P.M., Hopwood A.T. 1957. Insectivora and Chiroptera from the Miocene rocks of Kenya Colony. Fossil Mammals of Africa 13:1–35.

Cabrera A. 1925. Genera mammalium. Insectivora, Galeopithecia. Madrid: Museo Nacional de Ciencias Naturales.

Corbet G.B., Hanks J. 1968. A revision of the elephant-shrews, Family Macroscelididae. Bull. Br. Mus. Nat. Hist. (Zool.) 16:45–111.

Cote S., Werdelin L., Seiffert E.R., Barry J.C. 2007. Additional material of the enigmatic Early Miocene mammal *Kelba* and its relationship to the order Ptolemaiida. Proc. Natl. Acad. Sci. USA 104:5510–5515.

Cuvier G. 1798. Tableau élémentaire de l'histoire naturelle des animaux. Paris: Baudouin. p. 108.

De Witte G.F., Frechkop S. 1955. Sur une espèce encore inconnue de mammifère africain, *Potamogale ruwenzorii*, sp. n. B. Inst. Roy. Sci. Nat. Belg. 31:1.

Dobson G.E. 1882. A monograph of the Insectivora, systematic and anatomical. Part 1, including the families Erinaceidae, Centetidae and Solenodontidae. London: John van Voorst.

Donoghue M.J., Edwards E.J. 2014. Biome shifts and niche evolution in plants. Ann. Rev. Ecol. Evol. Syst. 45:547–572.

Douady C.J., Catzeflis F., Kao D.J., Springer M.S., Stanhope M.J. 2002. Molecular evidence for the monophyly of Tenrecidae (Mammalia) and the timing of the colonization of Madagascar by Malagasy tenrecs. Mol. Phylogenet. Evol. 22:357–363.

Douady C.J., Douzery E.J.P. 2003. Molecular estimation of eulipotyphlan divergence times and the evolution of "Insectivora". Mol. Phylogenet. Evol. 28:285–296.

Drummond A.J., Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7:214.

Du Chaillu P.B. 1860. The geographical features and natural history of a hitherto unexplored region of Western Africa. Proc. Roy. Geog. Soc. Lon. 5:361–363.

Eisenberg J.F. 1975. Phylogeny, behavior, and ecology in the Mammalia. In: Luckett W.P., Szalay F.S., editors. Phylogeny of the primates. New York: Plenum Press. p. 47–68.

Eisenberg J.F. 1981. The mammalian radiations: an analysis of trends in evolution, adaptation, and behavior. Chicago: University of Chicago Press.

FitzJohn R.G. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. Methods Ecol. Evol. 3:1084–1092.

Froriep L.F. 1806. C. Dumeril's analytische zoologie aus dem Französischen mit zusätzen von L.F. Froriep. Weimar, Germany: Landes-Industrie-Comptoir. p. 15.

Gaston K.J. 2000. Global patterns in biodiversity. Nature 405:220–227.

Gelman A., Rubin D. B. 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 457–472.

Genest H., Petter F. 1975. Part 1.1, Family Tenrecidae. In: Meester J., Setzer H.W., editors. The mammals of Africa: an identification manual. Washington, D.C.: Smithsonian Institution Press. p. 7.

Gheerbrant E., Amaghzaz M., Bouya B., Goussard F., Letenneur C. 2014. *Ocepeia* (Middle Paleocene of Morocco): the oldest skull of an afrotherian mammal. PLoS One 9:e89739.

Gheerbrant E., Sudre J., Cappetta H. 1996. A Palaeocene proboscidean from Morocco. Nature 383:68–70.

Gheerbrant E., Sudre J., Larochene M., Moumni A. 2001. First ascertained African "condylarth" mammals (primitive ungulates: cf. Bulbulodentata and cf. Phenacodonta) from the earliest Ypresian of the Ouled Abdoun Basin, Morocco. J. Vert. Paleontol. 21:107–118.

Goldberg E.E., Lancaster L.T., Ree R.H. 2011. Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. Syst. Biol. 60:451–465.

Goodman S.M., Jenkins P.D. 1998. The insectivores of the Réserve Spéciale d'Anjanaharibe-Sud, Madagascar. In: Goodman S.M., editor. A floral and faunal inventory of the Réserve Spéciale d'Anjanaharibe-Sud, Madagascar: with reference to elevational variation. Fieldiana Zool. (ns) 90:139.

Goodman S.M., Raxworthy C.J., Maminirina C.P., Olson L.E. 2006. A new species of shrew tenrec (*Microgale jobihely*) from northern Madagascar. J. Zool. (Lond.) 270:384–398.

Goodman S.M., Soarimalala V. 2004. A new species of *Microgale* (Lipotyphla: Tenrecidae: Oryzorictinae) from the Forêt des Mikea of southwestern Madagascar. Proc. Biol. Soc. Wash. 117:251–265.

Goodman S.M., Soarimalala V., Raheriarisena M., Rakotondravony D. 2013. Small mammals or tenrecs (Tenrecidae) and rodents (Nesomyidae). In: Goodman S.M., Raherilalao M., editors. Atlas of selected land vertebrates of Madagascar. Antananarivo, Madagascar: Association Vahatra. p. 211–269.

Goodman S.M., Vasey N., Burney D.A. 2007. Description of a new species of subfossil shrew tenrec (Afrosoricida: Tenrecidae: *Microgale*) from cave deposits in southeastern Madagascar. Proc. Biol. Soc. Wash. 120:367.

Grandidier A. 1870. Description de quelques animaux nouveaux découverts à Madagascar en novembre 1869. Rev. Mag. Zool. Pur. App. 22:50.

Grandidier G. 1899. Description d'une nouvelle espèce d'insectivore provenant de Madagascar. Bull. Mus. Natl. Hist. Nat. 5:349.

Grandidier G. 1934. Deux nouveaux mammifère insectivores de Madagascar *Microgale drouhardi* et *M. parvula*. Bull. Mus. Natl. Hist. Nat. 2:474–476.

Grandidier G. 1937. Mammifères nouveaux de la région de Diego-Suarez (Madagascar). Bull. Mus. Natl. Hist. Nat. 9:353.

Gray J.E. 1821. On the natural arrangement of vertebrose animals. Lon. Med. Repos. 15:301.

Gray J.E. 1837. Description of some new or little known Mammalia, principally in the British Museum Collection. Mag. Nat. Hist. (ns) 1:577–587.

Günther A. 1875. Notice of two new species of mammals (*Propithecus* and *Hemicentetes*) from Madagascar. Ann. Mag. Nat. Hist. 4 :125.

Guth C., Heim de Balsac H., Lamotte M. 1959. Recherches sur la morphologie de *Micropotamogale lamottei* et l'évolution des Potamogalinae: écologie, denture, anatomie crânienne. Mammalia 23:423–447.

Guth C., Heim de Balsac H., Lamotte M. 1960. Recherches sur la morphologie de *Micropotamogale lamottei* et l'évolution des Potamogalinae: rachis, viscères, position systématique. Mammalia 24:190–217.

Harper G., Steininger M., Tucker C., Juhn D., Hawkins F. 2007. Fifty years of deforestation and forest fragmentation in Madagascar. *Environ. Conserv*. 34:325–333.

Hartenberger J.L. 1986. Hypothèse paléontologique sur l'origine des Macroscelidea (Mammalia). C. R. Acad. Sci. II. 302:247–249.

Heim de Balsac H. 1954. Un genre inédit et inattendu de mammifère (Insectivore Tenrecidae) d'Afrique Occidentale. C. R. Acad. Sci. Paris 239:102–103.

Heled J., Drummond A.J. 2010. Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. 27:570–580.

Hillis D.M., Heath T.A., St. John K. 2005. Analysis and visualization of tree space. Syst. Biol. 54:471–482.

Huelsenbeck J., Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Jacobs L.L., Anyonge W., Barry J.C. 1987. A giant tenrecid from the Miocene of Kenya. J. Mammal. 68:10–16.

Jenkins P.D. 1992. Description of a new species of *Microgale* (Insectivora: Tenrecidae) from eastern Madagascar. Bull. Brit. Mus. Nat. Hist. Zool. 58:53–59.

Jenkins P.D. 1993. A new species of *Microgale* (Insectivora, Tenrecidae) from eastern Madagascar with an unusual dentition. Am. Mus. Nov. 3067:1–11.

Jenkins P.D. 2003. *Microgale*, shrew tenrecs. In: Goodman S.M., Benstead J.P., editors. The natural history of Madagascar. Chicago: University of Chicago Press. p. 1273–1278.

Jenkins P.D., Goodman S.M. 1999. A new species of *Microgale* (Lipotyphla, Tenrecidae) from isolated forest in southwestern Madagascar. Bull. Nat. Hist. Mus. Zool. 65:155–164.

Jenkins P.D., Goodman S.M., Raxworthy C.J. 1996. The shrew tenrecs (*Microgale*) (Insectivora: Tenrecidae) of the Réserve Naturelle Intégrale d'Andringitra, Madagascar. In: Goodman S.M., editor. A floral and faunal inventory of the eastern slopes of the Réserve Naturelle Intégrale d'Andringitra, Madagascar: with reference to elevational variation. Fieldiana Zool. n.s. 85:191–217.

Jenkins P.D., Raxworthy C.J., Nussbaum R.A. 1997. A new species of *Microgale* (Insectivora: Tenrecidae), with comments on the status of four other taxa of shrew tenrecs. Bull. Nat. Hist. Mus. Zool. 63:1–12.

Knowles L.L., Carstens B.C. 2007. Delimiting species without monophyletic gene trees. Syst. Biol. 56:887–895.

Kubatko L.S., Degnan J.H. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. 56:17–24.

Kumazawa Y., Nishida M. 1993. Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. J. Mol. Evol. 37:380–398.

Lacépède B.G.E. 1799. Tableau des divisions, sous-divisions, ordres et genres des mammifères. Paris: Chez Plassan. p. 7.

Lanfear R., Calcott B., Ho S.Y., Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29:1695–1701.

Lovegrove B.G., Lobban, K.D., Levesque, D.L. 2014. Mammal survival at the Cretaceous-Palaeogene boundary: metabolic homeostasis in prolonged tropical hibernation in tenrecs. Proc. Roy. Soc. B 281:20141304.

MacPhee R.D.E. 1987. The shrew tenrecs of Madagascar: systematic revision and Holocene distribution of *Microgale* (Tenrecidae, Insectivora). Am. Mus. Novit. 2889:1–45.

Maddison D., Maddison W. 2005. MacClade v. 4.08. Sunderland, MA: Sinauer Assoc.

Maddison W.P., Midford P.E., Otto S.P. 2007. Estimating a binary character's effect on speciation and extinction. Syst. Biol. 56:701–710.

Major C.F. 1896. Diagnoses of new mammals from Madagascar. J. Nat. Hist. 18:318–321.

Malia M.J., Jr., Adkins R.M., Allard M.W. 2002. Molecular support for Afrotheria and the polyphyly of Lipotyphla based on analyses of the growth hormone receptor gene. Mol. Phylogenet. Evol. 24:91–101.

Martin W.C.L. 1838. On a new genus of insectivorous Mammalia. Proc. Zool. Soc. London 6:17.

Matzke N.J. 2013. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. Front. Biogeogr. 5:242–248.

Matzke N.J. 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. Syst. Biol. 63:951–970.

McDowell S.B., Jr. 1958. The Greater Antillean insectivores. Bull. Am. Mus. Nat. Hist. 115:113–214.

McKenna M.C., Bell S.K. 1997. Classification of mammals above the species level. New York: Columbia University Press.

Meredith R.W., Janečka J.E., Gatesy J., Ryder O.A., Fisher C.A., Teeling E.C., Goodbla A., Eizirik E., Simão T.L.L., Stadler T., Rabosky D.L., Honeycutt R.L., Flynn J.J., Ingram C.M., Steiner C., Williams T.L., Robinson T.J., Burk-Herrick A., Westerman M., Ayoub N.A., Springer M.S., Murphy W.J. 2011. Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. Science 334:521–524.

Milne-Edwards A., Grandidier A. 1872. Description d'un nouveau mammifère insectivore de Madagascar (*Geogale aurita*). Ann. Sci. Nat. Zool. 15:1–5.

Milne-Edwards A., Grandidier A. 1882. Description d'une nouvelle espèce d'insectivore de Madagascar. Le Naturaliste 4:349.

Mittermeier R.A., Louis E.E. Jr, Langrand O., Schwitzer C., Gauthier C.A., Rylands A.B., Rajaobelina S, Ratzimbazafy J, Rasoloarison R., Hawkins F., Roos C., Richardson M., Kappeler P. 2014. Lémuriens de Madagascar. Paris: Publications scientifiques du Muséum national d'Histoire naturelle.

Mivart S.G. 1871. On *Hemicentetes*, a new genus of Insectivora, with some additional remarks on the osteology of that order. Proc. Zool. Soc. Lon. 1871:58–79.

Morales J., Pickford M., Salesa M., Soria D. 2000. The systematic status of *Kelba*, Savage, 1965, *Kenyalutra*, Schmidt-Kittler, 1987 and *Ndamathaia*, Jacobs et al., 1987 (Viverridae, Mammalia) and a review of Early Miocene mongoose-like carnivores of Africa. Ann. Paleontol. 86:243–251.

Mouchaty S.K., Gullberg A., Janke A., Arnason U. 2000. Phylogenetic position of the tenrecs (Mammalia: Tenrecidae) of Madagascar based on analysis of the complete mitochondrial genome sequence of *Echinops telfairi*. Zool. Scr. 29:307–317.

Nei M., Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3:418–426.

O'Leary M.A., Bloch J.I., Flynn J.J., Gaudin T.J., Giallombardo A., Giannini N.P., Goldberg S.L., Kraatz B.P., Luo Z.-X., Meng J., Ni X., Novacek M.J., Perini, F.A., Randall I.Z.S., Rougier G.W., Sargis E.J., Velazco P.M., Weksler M., Wible J.R., Cirranello A.L. 2013. The placental mammal ancestor and the post–K-Pg radiation of placentals. Science 339:662–667.

Olson L.E. 1999. Systematics, evolution, and biogeography of Madagascar's tenrecs (Mammalia: Tenrecidae). Ph.D. dissertation. Chicago: University of Chicago.

Olson L.E. 2013. Tenrecs. Current Biology 23: R5–R8.

Olson L.E., Goodman S.M. 2003. Phylogeny of Madagascar's tenrecs (Lipotyphla, Tenrecidae). In: Goodman S.M., Benstead J.P., editors. The natural history of Madagascar. Chicago: University of Chicago Press. p. 1235–1242.

Olson L.E., Goodman S.M., Yoder A.D. 2004. Illumination of cryptic species boundaries in long-tailed shrew tenrecs (Mammalia: Tenrecidae; *Microgale*), with new insights into geographic variation and distributional constraints. Biol. J. Linn. Soc. 83:1–22.

Olson L.E., Rakotomalala Z., Hildebrandt K.B.P., Lanier H.C., Raxworthy C.J., Goodman S.M. 2009. Phylogeography of *Microgale brevicaudata* (Tenrecidae) and description of a new species from western Madagascar. J. Mammal. 90:1095–1110.

Olson L.E., Sargis E.J., Martin R. 2005. Intraordinal phylogenetics of treeshrews (Mammalia: Scandentia) based on evidence from the mitochondrial 12S rRNA gene. Mol. Phylogenet. Evol. 35:656–673.

Paradis E., Claude J., Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290.

Parham J.F., Donoghue P.C.J., Bell C.J., Calway T.D., Head J.J., Holroyd P.A., Inoue J.G., Irmis R.B., Joyce W.G., Ksepka D.T., Patané J.S.L., Smith N.D., Tarver J.E., Van Tuinen M., Yang Z., Angielczyk K.D., Greenwood J.M., Hipsley C.A., Jacobs L., Makovicky P.J., Müller J., Smith K.T., Theodor J.M., Warnock R.C.M., Benton M.J. 2012. Best practices for justifying fossil calibrations. Syst. Biol. 61:346–359.

Posada D., Crandall K.A. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818.

Poux C., Madsen O., Glos J., de Jong W.W., Vences M. 2008. Molecular phylogeny and divergence times of Malagasy tenrecs: influence of data partitioning and taxon sampling on dating analyses. BMC Evol. Biol. 8:102.

Poux C., Madsen O., Marquard E., Vieites D.R., de Jong W.W., Vences M. 2005. Asynchronous colonization of Madagascar by the four endemic clades of primates, tenrecs, carnivores, and rodents as inferred from nuclear genes. Syst. Biol. 54:719–730.

Rabosky D.L., Goldberg E.E. 2015. Model inadequacy and mistaken inferences of trait-dependent speciation. Syst. Biol. 64:340–355.

Rambaut A., Suchard M.A., Xie D., Drummond A.J. 2014. Tracer v1.6. Available from: URL http://beast.bio.ed.ac.uk/Tracer.

Ree R.H., Moore B.R., Webb C.O., Donoghue M.J. 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. Evol. 59:2299–2311.

Roca A.L., Bar-Gal G.K., Eizirik E., Helgen K.M., Maria R., Springer M.S., O'Brien S.J., Murphy W.J. 2004. Mesozoic origin for West Indian insectivores. Nature 429:649–651.

Salton J.A., Szalay F.S. 2004. The tarsal complex of Afro-Malagasy Tenrecoidea: a search for phylogenetically meaningful characters. J. Mamm. Evol. 11:73–104.

Samonds K.E., Godfrey L.R., Ali J.R., Goodman S.M., Vences M., Sutherland M.R., Irwin M.T., Krause D.W. 2013. Imperfect isolation: factors and filters shaping Madagascar's extant vertebrate fauna. PLoS ONE 8:e62086.

Savage R.J.G. 1965. Fossil mammals of Africa: 19, The Miocene Carnivora of East Africa. Bull. Br. Mus. Nat. Hist. (Geol.) 10:239–316.

Schreber J.C.D. 1777. Die Säugthiere in Abbildungen nach der Natur mit Beschreibungen. Erlangen, Germany: Wolfgang Walther. p. 583–584.

Seiffert E.R., Simons E.L. 2000. *Widanelfarasia*, a diminutive placental from the late Eocene of Egypt. Proc. Natl. Acad. Sci. USA 97:2646–2651.

Seiffert E.R., Simons E.L., Ryan T.M., Bown T.M., Attia Y. 2007. New remains of Eocene and Oligocene Afrosoricida (Afrotheria) from Egypt, with implications for the origin(s) of afrosoricid zalambdodonty. J. Vert. Paleont. 27:963–972.

Simpson G.G. 1945. The principles of classification and the classification of mammals. Bull. Am. Mus. Nat. Hist. 85:1–350.

Soarimalala V., Goodman S.M. 2011. Les petits mammifères de Madagascar. Antananarivo, Madagascar: Association Vahatra.

Springer M.S., Douzery E. 1996. Secondary structure and patterns of evolution among mammalian mitochondrial 12S rRNA molecules. J. Mol. Evol. 43:357–373.

Stanhope M.J., Waddell V.G., Madsen O., De Jong W., Hedges S.B., Cleven G.C., Kao D., Springer M.S. 1998. Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. Proc. Nat. Acad. Sci. USA 95:9967–9972.

Stephenson P.J. 2003. *Geogale aurita*, large-eared tenrec. In: Goodman S.M., Benstead J.P., editors. The natural history of Madagascar. Chicago: University of Chicago Press. p. 1265–1267.

Sukumaran J., Holder M.T. 2010. DendroPy: a Python library for phylogenetic computing. Bioinformatics 26:1569–1571.

Tamura K., Stecher G., Peterson D., Filipski A., Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30:2725–2729.

Thomas O. 1882. Description of a new genus and two new species of Insectivora from Madagascar. J. Linn. Soc. Lon. 16:319–320.

Thomas O. 1918. On the arrangement of the small Tenrecidae hitherto referred to *Oryzorictes* and *Microgale*. Ann. Mag. Nat. Hist. 9:302–307.

Thomas O. 1926. On some small mammals from Madagascar. J. Nat. Hist. 17:250–252.

Trouessart E.-L. 1881 [for 1879]. Catalogue des mammifères vivants et fossiles. Revue et Magasin de Zoologie, sér. 3, 7:219–285.

Van Valen L. 1967. New Paleocene insectivores and insectivore classification. Bull. Am. Mus. Nat. Hist. 135:217–284.

Weisrock D.W. 2012. Concordance analysis in mitogenomic phylogenetics. Mol. Phylogenet. Evol. 65:194–202.

Wells N.A. 2003. Some hypotheses on the Mesozoic and Cenozoic paleoenvironmental history of Madagascar. In: Goodman S.M., Benstead J.P., editors. The natural history of Madagascar. Chicago: University of Chicago Press. p. 16–33.

Yoder A.D., Nowak M.D. 2006. Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. Annu. Rev. Ecol. Syst. 37:405–431.

Zwickl D.J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin.

1.8 Tables

Table 1.1. Fossil taxa used to calibrate the divergence time analysis (BEAST). Analytical parameters for hard-minimum and soft-maximum calibrations are provided. "Node" refers to the labeled, fossil-calibrated nodes on Figure 1.4.

Node	Taxon	Reference(s)	Age (MY)	Node Placement	Lognormal Distribution
A	†*Ocepeia*	Gheerbrant et al., 2001; Classification by Gheerbrant et al. (2014)	59.2-61.6	Afrotheria	offset=59.2, log(mean)=3, log(SD)=0.5
B	†*Phosphatherium*	Gheerbrant et al., 1996	48.6-55.8	Paenungulata	offset=48.6, log(mean)=2.5, log(SD)=0.5
C	†*Chambius*	Hartenberger, 1986; Cladistic analysis by Benoit et al. (2013)	40.4-55.8	Macroscelidea-Afrosoricida	offset=40.4, log(mean)=2.5, log(SD)=0.5
D	†*Miorhynchocyon*	Butler, 1984; Classification by McKenna & Bell (1997)	20.0-22.4	Macroscelidea	offset=20.0, log(mean)=2.5, log(SD)=0.5
E	†*Widanelfarasia*	Seiffert & Simons, 2000; Phylogenetic analysis by Seiffert et al. (2007)	33.9-37.8	Afrosoricida	offset=33.9, log(mean)=2.5, log(SD)=0.5
F	†*Parageogale*	Butler, 1984; Classification by McKenna & Bell (1997)	11.61-23.03	Geogalinae-Oryzorictinae	offset=11.608, log(mean)=2.0, log(SD)=0.5

Table 1.2. Characteristics of the four sequence data partitions estimated by PartitionFinder. Ambiguously aligned positions were removed from all analyses and are not included in these calculations. Calculations do not include the sequences from taxa outside Afrosoricida that were used for divergence dating.

Partition	Model	Genes (Codons)	Base Pairs	% Missing Data	Parsimony Informative Sites
1	HKY+I+G*	12S, tRNA-Val, ND2 (1,2)	1542	1.89%	536
2	GTR+I+G	ND2 (3)	348	3.28%	329
3	GTR+I+G	A2AB (1,2), AQP2 (1,2), AR (1,2), BDNF (1,2,3), GHR (1,2), RAG1 (1,2), vWF (1,2)	4516	4.29%	571
4	GTR+G	A2AB (3), AQP2 (3), AR (3), BRCA1 (1,2,3), GHR (3), RAG1 (3), vWF (3)	2974	4.74%	1493
		Total	9380	4.00%	2929

*In MrBayes analyses, the doublet model was assigned to all 12-S and tRNA-Val pairing stem regions.

Table 1.3. Mean ages and 95% highest posterior density ranges (HPD) for major afrosoricidan lineage splits.

Divergence event	**Mean Age (MY)**	**95% HPD (MY)**
Tenrecs - golden moles	58.44	50.83 - 66.19
African otter shrews - Malagasy tenrecs	48.33	40.73 - 55.61
Tenrecinae - Remaining Malagasy tenrecs	35.59	29.57 - 41.75
Geogalinae - Oryzorictinae	30.10	24.49 - 35.48
Oryzorictes - Remaining oryzorictines	27.81	22.68 - 33.22
(*M. dobsoni* + *M. talazaci*) - (*Microgale* + *Limnogale*)	19.35	15.85 - 23.12
Tenrec - *Hemicentetes*	16.10	10.8 - 21.72
Echinops - *Setifer*	10.21	5.24 - 15.89
Limnogale mergulus - *M. parvula*	14.96	11.57 - 18.07
Geogale - *Geogale*	9.19	4.91 - 13.59

Table 1.4. Comparison of alternative GeoSSE submodels using the R package diversitree. Parameters are speciation (s), extinction (x), and dispersal (d) for species in humid habitats (H), dry habitats (D), and both (HD). d.f. = degrees of freedom.

Model	d.f.	-lnL	ΔAIC	AIC weight	sA	sB	sAB	xA	xB	dA	dB
xH=0, xD=0, dH=dD	4	-129.97	0.00	0.436	0.070	0.024	0.217	-	-	0.018	-
xH=0, xD=0, sH=sD, dH=dD	3	-131.88	1.81	0.176	0.057	-	0.225	-	-	0.018	-
xH=0, xD=0	5	-129.95	1.94	0.165	0.070	0.024	0.234	-	-	0.017	0.023
xH=0, xD=0, sH=sD	4	-131.80	3.66	0.070	0.057	-	0.240	-	-	0.016	0.025
dH=dD	6	-129.95	3.94	0.060	0.070	0.024	0.217	1.60E-07	2.89E-06	0.018	-
xH=xD	6	-129.97	4.00	0.059	0.070	0.024	0.234	1.00E-06	-	0.017	0.023
Full	7	-129.95	5.94	0.022	0.070	0.024	0.234	2.33E-06	5.17E-06	0.017	0.023
sH=sD	6	-131.69	7.43	0.011	0.060	-	0.237	1.00E-08	2.10E-02	0.019	0.021
sHD=0	6	-134.53	13.11	0.001	0.103	0.032	-	0.052	0.019	0.017	0.017

Table 1.5. Revised taxonomy of Tenrecomorpha.

Tenrecomorpha Butler 1972:261
- **Potamogalidae** Allman, 1865:467
  - *Micropotamogale* Heim de Balsac, 1954:102
    - *M. lamottei* Heim de Balsac, 1954:103
    - *M. ruwenzorii* (de Witte and Frechkop, 1955:1)
  - [3]*Potamogale* du Chaillu, 1860:363
    - *P. velox* (du Chaillu, 1860:361)
- **Tenrecidae** Gray, 1821:301
  - **Geogalinae** Trouessart, 1879:275
    - *Geogale* Milne-Edwards and G. Grandidier, 1872:1.
      - *G. aurita* Milne-Edwards and G. Grandidier, 1872:1
  - **Oryzorictinae** Dobson, 1882b:71
    - *Microgale* Thomas, 1882:319
      - *M. brevicaudata* G. Grandidier, 1899:349
      - *M. cowani* Thomas, 1882:320
      - *M. drouhardi* G. Grandidier, 1934:474
      - *M. dryas* Jenkins, 1992:53
      - *M. fotsifotsy* Jenkins, Raxworthy, and Nussbaum, 1997:2
      - *M. gracilis* (Major, 1896a:321)
      - *M. grandidieri* Olson, Rakotomalala, Hildebrandt, Lanier, Raxworthy, and Goodman, 2009:1097
      - *M. gymnorhyncha* Jenkins, Goodman, and Raxworthy, 1996:211
      - *M. jenkinsae* Goodman and Soarimalala, 2004:251
      - *M. jobihely* Goodman, Raxworthy, Maminirina, and Olson, 2006:384
      - *M. longicaudata* Thomas, 1882:320
      - †*M. macpheei* Goodman, Vasey, and Burney, 2007:367
      - *M. majori* Thomas, 1918:304
      - [1]*M. mergulus* (Major, 1896a:318)
      - *M. monticola* Goodman and Jenkins, 1998:149
      - *M. nasoloi* Jenkins and Goodman, 1999:156
      - *M. parvula* G. Grandidier, 1934:476
      - *M. principula* Thomas, 1926:250
      - *M. pusilla* Major, 1896b:462
      - *M. soricoides* Jenkins, 1993:2
      - *M. taiva* Major, 1896b:461
      - *M. thomasi* Major, 1896a:318
    - *Nesogale* Thomas 1918:303
      - [2]*Nesogale dobsoni* (Thomas, 1884:337)
      - [2]*Nesogale talazaci* (Major, 1896a:318)
    - *Oryzorictes* A. Grandidier, 1870:50
      - *O. hova* A. Grandidier, 1870:50
      - *O. tetradactylus* Milne-Edwards and G. Grandidier, 1882:55
  - **Tenrecinae** Gray, 1821:301
    - *Echinops* Martin, 1838:17
      - *E. telfairi* Martin, 1838:17
    - *Hemicentetes* Mivart, 1871:58
      - *H. nigriceps* Günther, 1875:125
      - *H. semispinosus* (G. Cuvier, 1798:108)
    - *Setifer* Froriep, 1806:15
      - *S. setosus* (Schreber, 1777:583)
    - *Tenrec* Lacépède, 1799:7
      - *T. ecaudatus* (Schreber, 1777:584).

[1]Formerly *Limnogale*. [2]Formerly *Microgale*. [3]Du Chaillu (1860) provisionally named *Potamogale* as the eventual genus to contain *velox* but, having only a skin on which to base his description, formally described the new species as *Cynogale velox*; *Cynogale* Gray 1837 is a genus of Carnivora.

1.9 Figures

Figure 1.1. Images of tenrecs highlighting a number of ecomorphological, physiological, and behavioral specializations. a) Long-tailed shrew tenrec, *Microgale longicaudata* (photo by L.E. Olson). b) Mole-like rice tenrec, *Oryzorictes hova* (photo by L.E. Olson). c) Aquatic or web-footed tenrec, *Limnogale mergulus* (photo by S. Zak). d) Lowland streaked tenrec, *Hemicentetes semispinosus* (photo by J.L. Fiely). e) Common or tailless tenrec, *Tenrec ecaudatus* (photo by L.E. Olson). f) Lesser hedgehog tenrec, *Echinops telfairi* (photo by L.E. Olson). g) Large-eared tenrec, *Geogale aurita* (photo by L.E. Olson).

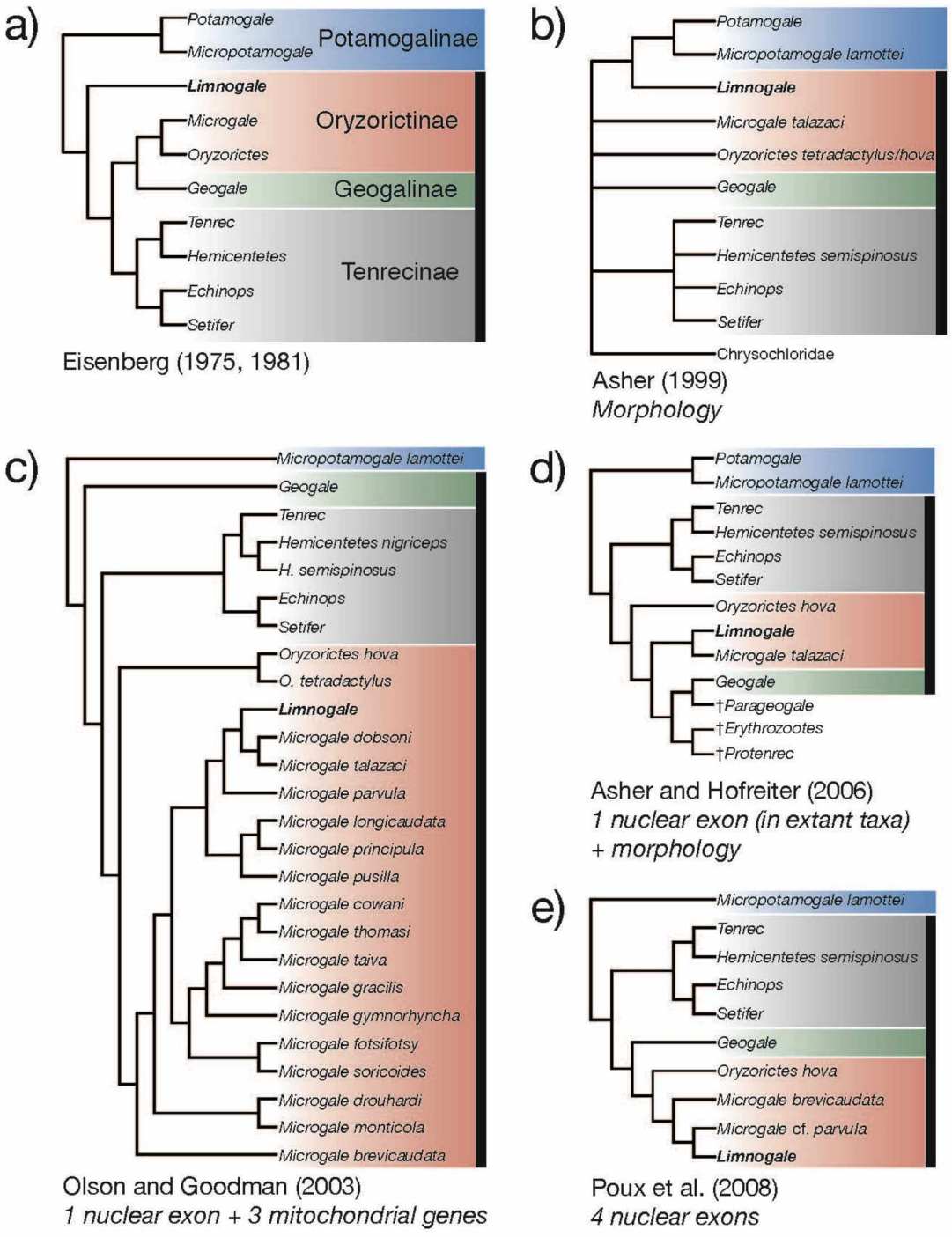


Figure 1.2. Previous phylogenetic hypotheses of Tenrecidae *sensu lato*. Traditionally recognized subfamilies are color coded for extant taxa only. Eisenberg's (1975, 1981) hypothesis was speculative and not based on any formal analysis. Trees b-e were rooted using outgroups (not shown). For monotypic genera only the genus name is presented. The black vertical line denotes Malagasy taxa.

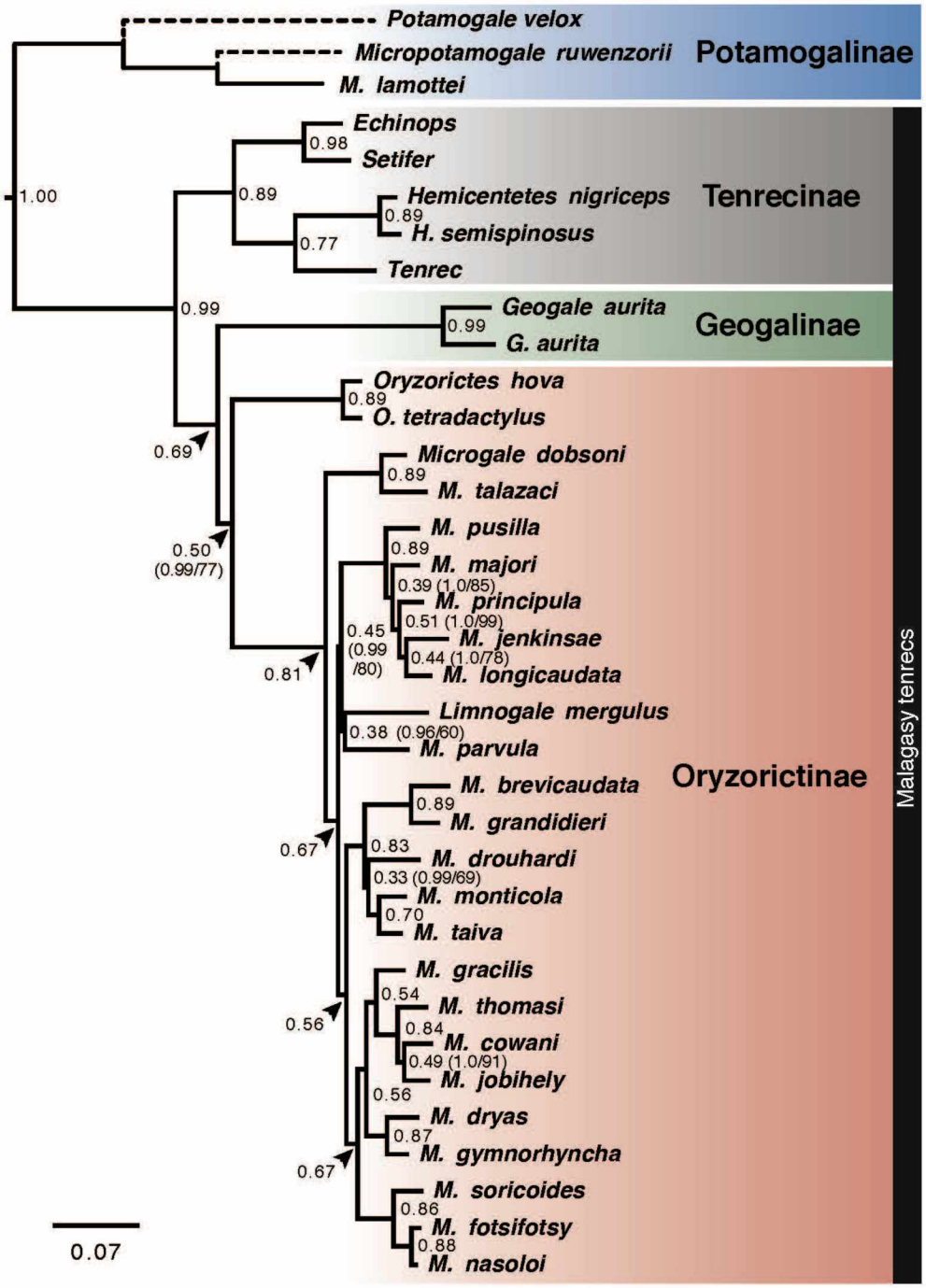


Figure 1.3. Phylogenetic relationships among Malagasy tenrecs and African otter shrews as determined from the BUCKy concordance analysis and the ML and BI concatenated analyses, which recovered identical topologies. Concordance factor values are shown at each node, and Bayesian posterior probabilities/ML bootstrap values are given in parentheses. Bayesian posterior probabilities >0.99 and ML bootstrap values >99 are not labeled. Branch lengths were produced by the BI concatenated analysis. Dashed lines indicate taxa for which only mitochondrial sequence data were collected.

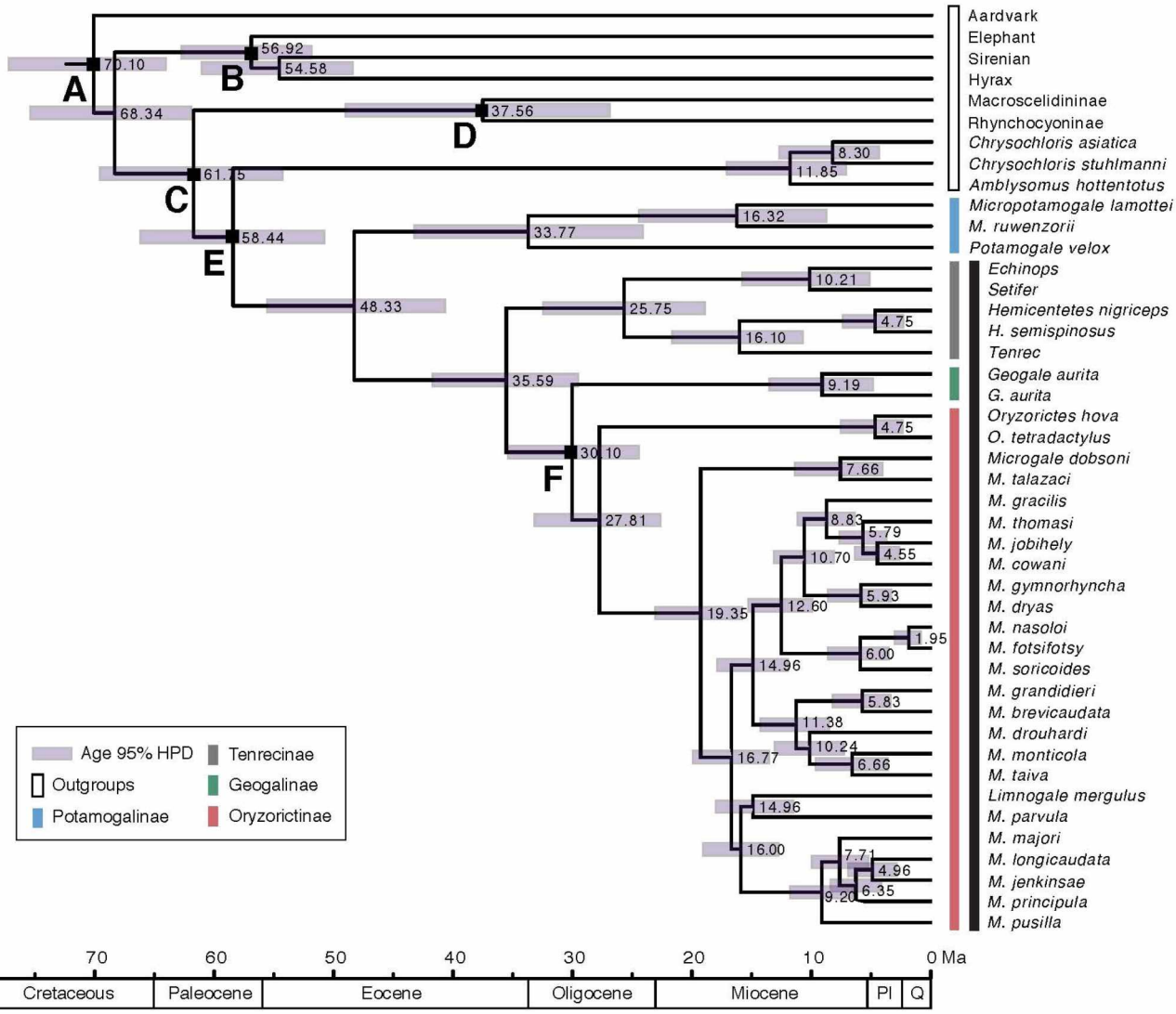


Figure 1.4. Divergence time chronogram produced using BEAST. Fossil calibration points are indicated by letters A-F next to black squares. Mean divergence time estimates (in millions of years) are adjacent to their respective nodes. Purple nodal bars correspond to the 95% highest posterior density regions. "Pl" = Pliocene, "Q" = Quaternary. Subfamilies and Malagasy taxa colored as in Figure 2. Tree was rooted using an armadillo (*Dasypus novemcinctus*) as the outgroup.

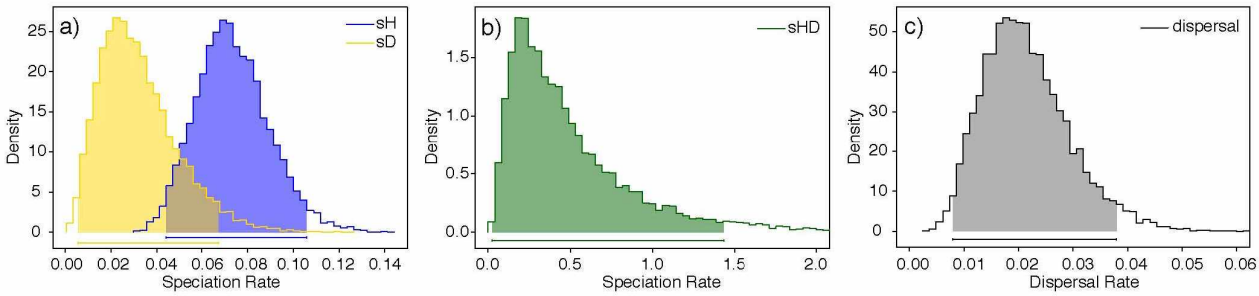


Figure 1.5. Posterior probability distributions of GeoSSE parameters for Madagascar's tenrecs. Rates of speciation were estimated for a) humid (H; blue) and dry (D; yellow) habitats, and for b) both habitats (HD; green). The top-ranking model estimated c) a single rate of dispersal for all habitats. Shaded regions in all plots show the 95% credible interval. The top-ranking model did not parameterize extinction. Note the difference in scale for all plots.

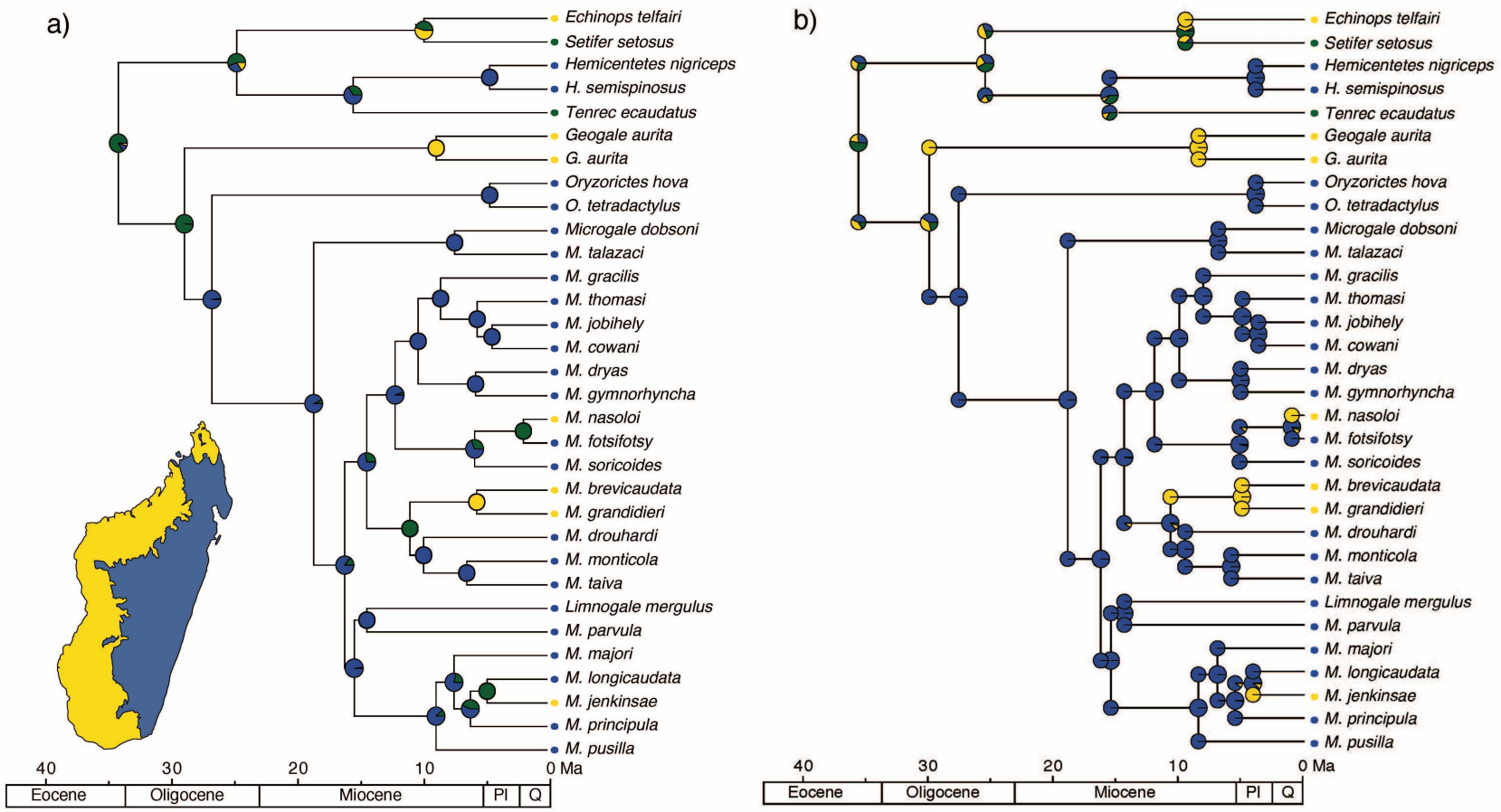


Figure 1.6. Ancestral character state reconstructions of the habitat associations in Malagasy tenrecs using a) the GeoSSE model and b) BioGeoBEARs' DEC+j model. Colors correspond to three habitat states: humid (green), dry (yellow), and both/eurytopic (blue). Pie charts show the relative likelihoods of each ancestral habitat. Geological epochs are shown below the timescale. The GeoSSE models reconstructs a single state for ancestral nodes, while BioGeoBEARs reconstructs ancestral nodes as a triplet of parent and daughter states.

1.10 Supplementary Tables

Supplementary Table 1.1 Characteristics of each gene examined in this study. Datasets do not include the non-afrosoricidan sequence data used for divergence dating. Parentheses show alignment statistics with ambiguously aligned characters removed for analyses.

	Mitochondrial Genes		Nuclear Genes		
Species	**12S+val**	**ND2**	**A2AB**	**AQP2**	**AR**
Amblysomus hottentotus	**M95108**	KX015410 (FMNH 165582)	**Y12526**	**Y15952**	**AJ893562**
Chrysochloris asiatica	KX015138 (FMNH 165435)	KX015411 (FMNH 165435)	**JN413821**	KX015200 (FMNH 165435)	-
Chrysochloris stuhlmanni	KX015139 (FMNH 137590)	KX015412 (FMNH 137590)	KX015173 (FMNH 137590)	KX015201 (FMNH 137590)	KX015267 (FMNH 137590)
Micropotamogale lamottei	**AF390539**	KX015413 (BMNH 1973.170)	**AJ251107**	**AJ251106**	**AJ893571**
Micropotamogale ruwenzorii	KX015140 (MRAC 31042)	KX015414 (MRAC 31042)	-	-	-
Potamogale velox	KX015141 (USNM 266897)	KX015415 (USNM 266897)	-	-	-
Echinops telfairi	KX015142 (FMNH 159739)	KX015416 (FMNH 159739)	**Y17692**	KX015202 (FMNH 159739)	**AJ893565**
Geogale aurita (1)	KX015143 (FMNH 159732)	KX015417 (FMNH 159732)	KX015174 (FMNH 159732)	KX015207 (FMNH 159732)	KX015269 (FMNH 159732)
Geogale aurita (2)	-	KX015418 (MVZ 220648)	**AM905342**	KX015208 (MVZ 220648)	**AM905338**
Hemicentetes nigriceps	KX015144 (UADBA 10815)	KX015419 (UADBA 10815)	KX015199 (UADBA 10815)	KX015204 (UADBA 10815)	KX015271 (UADBA 10815)
Hemicentetes semispinosus	KX015145 (UADBA 10596)	KX015420 (UADBA 10596)	**AJ891065**	KX015205 (UADBA 10596)	**AJ893567**
Limnogale mergulus	KX015146 (FMNH 165440)	KX015421 (FMNH 165440)	**AJ891069**	KX015211 (FMNH 165440)	**AJ893570**
Microgale brevicaudata	KX015147 (FMNH 156310)	KX015422 (FMNH 156310)	**AJ891072**	KX015212 (FMNH 156310)	KX015282 (FMNH 156310)
Microgale cowani	KX015148 (FMNH 156560)	KX015423 (FMNH 156560)	KX015195 (FMNH 156560)	KX015213 (FMNH 156560)	KX015280 (FMNH 156560)
Microgale dobsoni	KX015149 (FMNH 161723)	KX015424 (FMNH 161723)	KX015182 (FMNH 156202)	KX015214 (FMNH 156413)	KX015273 (FMNH 161723)
Microgale drouhardi	KX015150 (FMNH 156324)	KX015425 (FMNH 156324)	KX015178 (FMNH 156324)	KX015215 (FMNH 156324)	KX015283 (FMNH 156324)
Microgale dryas	KX015151 (UADBA 47148)	KX015426 (FMNH 176395)	KX015197 (FMNH 176381)	KX015216 (FMNH 176395)	KX015285 (FMNH 176395)
Microgale fotsifotsy	KX015152 (FMNH 156312)	KX015427 (FMNH 156312)	KX015189 (FMNH 156312)	KX015217 (FMNH 156312)	KX015287 (FMNH 156312)
Microgale gracilis	KX015153 (UADBA 32501)	KX015428 (UADBA 32501)	KX015192 (UADBA 32501)	KX015218 (UADBA 32501)	KX015278 (UADBA 32501)
Microgale grandidieri	KX015154 (FMNH 169678)	KX015429 (FMNH 169678)	KX015177 (FMNH 173231)	KX015219 (FMNH 169678)	KX015268 (FMNH 169678)
Microgale gymnorhyncha	KX015155 (FMNH 159666)	KX015430 (FMNH 159666)	KX015196 (FMNH 159666)	KX015220 (FMNH 159666)	KX015286 (FMNH 159666)
Microgale jenkinsae	KX015156 (FMNH 176154)	KX015431 (FMNH 176154)	KX015184 (FMNH 176215)	KX015221 (FMNH 176154)	KX015290 (FMNH 176154)
Microgale jobihely	KX015157 (AMNH 274988)	KX015432 (AMNH 274985)	KX015193 (AMNH 274984)	KX015222 (AMNH 274988)	KX015281 (AMNH 274988)
Microgale longicaudata	KX015158 (FMNH 156317)	KX015433 (FMNH 156317)	KX015183 (FMNH 156317)	KX015223 (FMNH 156317)	KX015291 (FMNH 156317)
Microgale majori	KX015159 (FMNH 161735)	KX015434 (FMNH 151632)	KX015187 (FMNH 156579)	KX015224 (FMNH 161735)	KX015272 (FMNH 161735)
Microgale monticola	KX015160 (FMNH 159678)	KX015435 (FMNH 159678)	KX015179 (FMNH 159678)	KX015225 (FMNH 159678)	KX015275 (FMNH 159678)
Microgale nasoloi	KX015161 (FMNH 161575)	KX015436 (FMNH 161575)	KX015190 (FMNH 161575)	KX015226 (FMNH 161575)	KX015277 (FMNH 161575)
Microgale parvula	KX015162 (FMNH 156329)	KX015437 (FMNH 156329)	KX015191 (FMNH 156329)	KX015227 (FMNH 156329)	KX015289 (FMNH 156329)
Microgale principula	KX015163 (FMNH 156591)	KX015438 (FMNH 156591)	KX015185 (FMNH 156591)	KX015228 (FMNH 156591)	KX015276 (FMNH 156591)
Microgale pusilla	KX015164 (FMNH 165489)	KX015439 (FMNH 165489)	KX015186 (FMNH 165489)	KX015229 (FMNH 165489)	KX015292 (FMNH 165489)
Microgale soricoides	KX015165 (FMNH 156594)	KX015440 (FMNH 156594)	KX015188 (FMNH 156594)	KX015230 (FMNH 156594)	KX015288 (FMNH 156594)
Microgale taiva	KX015166 (FMNH 162011)	KX015441 (FMNH 162011)	KX015180 (FMNH 162011)	KX015231 (FMNH 162011)	KX015284 (FMNH 162011)
Microgale talazaci	KX015167 (FMNH 156332)	KX015442 (FMNH 156332)	KX015181 (FMNH 156332)	KX015232 (FMNH 156332)	KX015274 (FMNH 156332)
Microgale thomasi	KX015168 (FMNH 156599)	KX015443 (FMNH 156599)	KX015194 (FMNH 156599)	KX015233 (FMNH 156599)	KX015279 (FMNH 156599)
Oryzorictes hova	KX015169 (FMNH 159466)	KX015444 (FMNH 159466)	KX015175 (FMNH 159466)	KX015209 (FMNH 159466)	**AJ893568**
Oryzorictes tetradactylus	KX015170 (FMNH 156226)	KX015445 (FMNH 156226)	KX015176 (FMNH 156226)	KX015210 (FMNH 156226)	KX015270 (FMNH 156226)
Setifer setosus	KX015171 (FMNH 159731)	KX015446 (FMNH 159731)	KX015198 (FMNH 159731)	KX015203 (FMNH 159731)	**AJ893566**
Tenrec ecaudatus	KX015172 (FMNH 156657)	KX015447 (FMNH 156657)	**AJ251108**	KX015206 (FMNH 156657)	**AJ893564**

Supplementary Table 1.1 (continued).

	Nuclear Genes				
Species	**BDNF**	**BRCA1**	**GHR**	**RAG1**	**vWF**
Amblysomus hottentotus	KX015234 (FMNH 165582)	**AF284027**	**JN414680**	KX015347 (FMNH 165582)	**U97534**
Chrysochloris asiatica	KX015235 (FMNH 165435)	**JN414131**	**JN414681**	KX015348 (FMNH 165435)	KX015380 (FMNH 165435)
Chrysochloris stuhlmanni	KX015236 (FMNH 137590)	KX015293 (FMNH 137590)	KX015324 (FMNH 137590)	KX015349 (FMNH 137590)	KX015381 (FMNH 137590)
Micropotamogale lamottei	**JQ073056**	**AY121759**	**DQ202290**	**JQ073179**	**AF390538**
Micropotamogale ruwenzorii	-	-	-	-	-
Potamogale velox	-	-	**DQ202291**	-	-
Echinops telfairi	**AY059686**	**AF284025**	**AF392889**	**JQ073181**	**AF076478**
Geogale aurita (1)	KX015242 (FMNH 159732)	KX015297 (FMNH 159732)	**DQ202287**	KX015354 (FMNH 159732)	KX015382 (FMNH 159732)
Geogale aurita (2)	KX015243 (MVZ 220648)	KX015298 (MVZ 220648)	**AM905347**	KX015355 (MVZ 220648)	**AM905352**
Hemicentetes nigriceps	KX015238 (UADBA 10815)	KX015295 (UADBA 10815)	KX015325 (UADBA 10815)	KX015351 (UADBA 10815)	KX015383 (UADBA 10815)
Hemicentetes semispinosus	KX015239 (UADBA 10596)	KX015296 (UADBA 10596)	**DQ202288**	KX015352 (UADBA 10596)	**AJ891093**
Limnogale mergulus	KX015244 (FMNH 165440)	KX015301 (FMNH 165440)	**DQ202289**	**JQ073180**	**AJ891096**
Microgale brevicaudata	KX015245 (FMNH 156310)	KX015302 (FMNH 156310)	**AM905345**	KX015358 (FMNH 156310)	KX015384 (FMNH 156310)
Microgale cowani	KX015246 (FMNH 156560)	KX015303 (FMNH 156560)	KX015327 (FMNH 156560)	KX015359 (FMNH 156560)	KX015385 (FMNH 156560)
Microgale dobsoni	KX015247 (FMNH 161723)	KX015304 (FMNH 161723)	KX015328 (FMNH 161723)	KX015360 (FMNH 161723)	KX015386 (FMNH 156413)
Microgale drouhardi	KX015248 (FMNH 156324)	KX015305 (FMNH 156324)	KX015329 (FMNH 156324)	KX015361 (FMNH 156324)	KX015387 (FMNH 156324)
Microgale dryas	KX015249 (FMNH 176395)	KX015306 (FMNH 176395)	KX015330 (FMNH 176395)	KX015362 (FMNH 176395)	KX015388 (FMNH 176381)
Microgale fotsifotsy	KX015250 (FMNH 156312)	KX015307 (FMNH 156312)	KX015331 (FMNH 156312)	KX015363 (FMNH 156312)	KX015389 (FMNH 156312)
Microgale gracilis	KX015251 (UADBA 32501)	KX015308 (UADBA 32501)	KX015332 (UADBA 32501)	KX015364 (UADBA 32501)	KX015390 (UADBA 32501)
Microgale grandidieri	KX015252 (FMNH 169678)	KX015309 (FMNH 169678)	KX015333 (FMNH 169678)	KX015365 (FMNH 169678)	KX015391 (FMNH 173239)
Microgale gymnorhyncha	KX015253 (FMNH 159666)	KX015310 (FMNH 159666)	KX015334 (FMNH 159666)	KX015366 (FMNH 159666)	KX015392 (FMNH 159666)
Microgale jenkinsae	KX015254 (FMNH 176154)	KX015311 (FMNH 176154)	KX015335 (FMNH 176154)	KX015367 (FMNH 176154)	KX015393 (FMNH 176215)
Microgale jobihely	KX015255 (AMNH 274988)	KX015312 (AMNH 274988)	KX015336 (AMNH 274988)	KX015368 (AMNH 274988)	KX015394 (AMNH 274988)
Microgale longicaudata	KX015256 (FMNH 156317)	KX015313 (FMNH 156317)	KX015337 (FMNH 156317)	KX015369 (FMNH 156317)	KX015395 (FMNH 156317)
Microgale majori	KX015257 (FMNH 151632)	KX015314 (FMNH 151632)	KX015338 (FMNH 151632)	KX015370 (FMNH 151632)	KX015396 (FMNH 161735)
Microgale monticola	KX015258 (FMNH 159678)	KX015315 (FMNH 159678)	KX015339 (FMNH 159678)	KX015371 (FMNH 159678)	KX015397 (FMNH 159678)
Microgale nasoloi	KX015259 (FMNH 161575)	KX015316 (FMNH 161575)	KX015340 (FMNH 161575)	KX015372 (FMNH 161575)	KX015398 (FMNH 161575)
Microgale parvula	KX015260 (FMNH 156329)	KX015317 (FMNH 156329)	KX015341 (FMNH 156329)	KX015373 (FMNH 156329)	KX015399 (FMNH 156329)
Microgale principula	KX015261 (FMNH 156591)	KX015318 (FMNH 156591)	KX015342 (FMNH 156591)	KX015374 (FMNH 156591)	KX015400 (FMNH 170764)
Microgale pusilla	KX015262 (FMNH 165489)	KX015319 (FMNH 165489)	KX015343 (FMNH 165489)	KX015375 (FMNH 165489)	KX015401 (FMNH 165489)
Microgale soricoides	KX015263 (FMNH 156594)	KX015320 (FMNH 156594)	KX015344 (FMNH 156594)	KX015376 (FMNH 156594)	KX015402 (FMNH 156594)
Microgale taiva	KX015264 (FMNH 162011)	KX015321 (FMNH 162011)	KX015345 (FMNH 162011)	KX015377 (FMNH 162011)	KX015403 (FMNH 162011)
Microgale talazaci	KX015265 (FMNH 156332)	KX015322 (FMNH 156332)	**AF392885**	KX015378 (FMNH 156332)	KX015404 (FMNH 156332)
Microgale thomasi	KX015266 (FMNH 156599)	KX015323 (FMNH 156599)	KX015346 (FMNH 156599)	KX015379 (FMNH 156599)	KX015405 (FMNH 156599)
Oryzorictes hova	**AY986750**	KX015299 (FMNH 159466)	**AF392886**	KX015356 (FMNH 159466)	KX015406 (FMNH 159466)
Oryzorictes tetradactylus	KX015241 (FMNH 156226)	KX015300 (FMNH 156226)	KX015326 (FMNH 156226)	KX015357 (FMNH 156226)	KX015407 (FMNH 156226)
Setifer setosus	KX015237 (FMNH 159731)	KX015294 (FMNH 159731)	**DQ202292**	KX015350 (FMNH 159731)	KX015408 (FMNH 159731)
Tenrec ecaudatus	KX015240 (FMNH 156657)	**AF284026**	**AF392890**	KX015353 (FMNH 156657)	KX015409 (FMNH 156657)

Supplementary Table 1.2. Primers used to amplify each gene for each species. Flanking forward and reverse primers are separated by a slash, and additional nested primers are listed in parentheses. Values in bold are GenBank accession numbers. Dashes indicate missing data. Primer sequences are listed in Supplementary Table 1.3.

Species	12S+val	ND2
Amblysomus hottentotus	**M95108**	1/2 (3TX, LOR2)
Chrysochloris asiatica	C/G (1, AF6, 2, LOR2)	1/2 (3TX, LOR2)
Chrysochloris stuhlmanni	C/2 (1, LOF1, LOR1)	1/2 (4)
Chrysospalax trevelyani	**AY249238**	-
Eremitalpa granti	**AM904729**	**AM904729**
Micropotamogale lamottei	**AF390539**	1/4
Micropotamogale ruwenzorii	C/G (1, LOF1, LOR2, 2)	1/2 (3TX, PF2, PotR, 4)
Potamogale velox	C/G (1, LOF1, 2, LOR1)	1/PotR (3TX, PoveF, MilaF, R1, LOR3, PoveR, LOR2, MilaR1, 4)
Echinops telfairi	C/G (1, LOF1, 2, LOR2)	1/2 (3TX, 4)
Geogale aurita (1)	C/G (1, LOF1, 2, LOR1, LOR2)	1/2 (3TX, NN, 2T, LOR2)
Geogale aurita (2)	-	F1AG/LOR2
Hemicentetes nigriceps	C/G (1, LOF1, LOR2)	1/2 (3TX, LOR2)
Hemicentetes semispinosus	C/G (1, LOF1, 2, LOR1)	1/2 (3TX, 4)
Limnogale mergulus	C/G (1, LOF1, 2, LOR2)	1/2 (3TX, LOR2)
Microgale brevicaudata	C/G (1, AF6, 2, LOR1)	1/2 (3TX, 4, LOR2)
Microgale cowani	C/G (1, LOF1, 2, LOR2)	1/2 (3TX, 4)
Microgale dobsoni	C/G (1, LOF1, 2, LOR1)	1/2 (3TX, LOR2)
Microgale drouhardi	C/G (1, LOF1, 2, LOR2)	NN/2 (3TX, LOR2)
Microgale dryas	C/G (AF4, AR3)	1/2T (3TX, LOR2)
Microgale fotsifotsy	C/G (1, LOF1, 2, LOR1, LOR2)	1/2 (3TX, 3, 4)
Microgale gracilis	C/G (AF6, LOR2)	1/2 (3TX, LOR2, 4)
Microgale grandidieri	C/G (TuF3, TuR3)	1/2 (BrF1, MR2, LOR2)
Microgale gymnorhyncha	C/G (1, LOF1, 2, LOR1)	1/2 (3TX, 3TY, 2T, MR1, LOR2)
Microgale jenkinsae	C/G (AF4, AR3)	1/2T (3TX, LOR2)
Microgale jobihely	C/G (TuF3, TuR3)	1/2 (LOF1, LOR2)
Microgale longicaudata	C/G (1, LOF1, LOR2)	1/2 (3TX, 4, LOR2)
Microgale majori	C/G (1,2)	1/2T (3TX, LOR2)
Microgale monticola	C/G (1, AF6, 2, LOR2)	1/2 (3TX, LOR2)
Microgale nasoloi	C/G (1, AF4, AF6, LOF1, AR6, AR3, LOR2)	1/2 (3TX, LOR2)
Microgale parvula	C/G (1, LOF1, 2, LOR2)	1/2 (3, 4)
Microgale principula	C/G (1, 2, LOR1, LOR2)	1/LOR2 (3TX, 3TY, 4)
Microgale pusilla	C/G (1, LOF1, 2, LOR2)	1/2 (3TX, LOR2)
Microgale soricoides	C/G (1, LOF1, 2, LOR2)	1/2 (3, 4, 3TY, MR3)
Microgale taiva	C/AR6 (1, AF4, 2, LOR2)	1/2 (3TX, 2T, LOR2)
Microgale talazaci	C/G (1, LOF1, 2, LOR2)	NN2/2 (3TX, LOR2)
Microgale thomasi	C/G (1, LOF1, 2, LOR2)	1/2 (3TX, LOR2)
Oryzorictes hova	C/G (1, LOF1, 2, LOR2)	1/2 (3TX, 4, LOR2)
Oryzorictes tetradactylus	C/G (1, 12G, LOR2)	1/2 (3TX, 4, LOR2)
Setifer setosus	C/G (1, LOF1, 2, LOR2)	1/2 (3TX, 2T, LOR2)
Tenrec ecaudatus	C/G (1, LOF1, 2, LOR2)	1/2 (3TX, LOR2)

Supplementary Table 1.2 (continued).

Species	A2AB	Aqp2	AR	BDNF
Amblysomus hottentotus	**Y12526**	**Y15952**	**AJ893562**	LOF1/LOR1
Chrysochloris asiatica	F3/R1 (F2, R3, R4)	LOF1/LOR1	LOF2/KER1	LOF1/LOR1
Chrysochloris stuhlmanni	F1/R1 (F2, R2)	LOF1/LOR1	LOF2/LOR1	LOF1/LOR1
Chrysospalax trevelyani	-	-	-	**AY986752**
Eremitalpa granti	-	-	-	-
Micropotamogale lamottei	**AJ251107**	**AJ251106**	**AJ893571**	**JQ073056**
Micropotamogale ruwenzorii	-	-	-	-
Potamogale velox	-	-	-	-
Echinops telfairi	**Y17692**	LOF1/LOR1	**AJ893565**	**AY059686**
Geogale aurita (1)	F1/R1 (F2, R3)	LOF1/LOR1	**AM905338**	LOF1/LOR1
Geogale aurita (2)	**AM905342**	KEF2/KER2	**AM905338**	KEF1/KER1
Hemicentetes nigriceps	F1/R1 (F2, R3)	LOF1/LOR1	LOF2/LOR1	LOF1/LOR1
Hemicentetes semispinosus	**AJ891065**	KEF1/LOR2	**AJ893567**	LOF1/LOR1
Limnogale mergulus	**AJ891069**	LOF2/LOR1	**AJ893570**	LOF1/LOR1
Microgale brevicaudata	**AJ891072**	LOF2/LOR1	LOF2/LOR1	LOF1/LOR1
Microgale cowani	F3/R1 (F2, R3)	LOF2/LOR1	LOF2/LOR1	LOF1/LOR1
Microgale dobsoni	F1/R1 (F2, R3)	LOF2/KER1	LOF2/KER1	LOF1/LOR1
Microgale drouhardi	F1/R1 (F2, R3)	KEF1/LOR2	LOF2/LOR1	LOF1/LOR1
Microgale dryas	F4/R6 (F2, R7)	KEF1/KER1	LOF2/LOR1	LOF1/LOR1
Microgale fotsifotsy	F1/R1 (F2, R3)	LOF2/LOR2	LOF2/LOR1	LOF1/LOR1
Microgale gracilis	F1/R1 (F2, R3)	LOF1/LOR1	LOF2/LOR1	LOF1/LOR1
Microgale grandidieri	F3/R6 (F2, R3)	KEF1/LOR2	LOF2/LOR1	LOF1/LOR1
Microgale gymnorhyncha	F1/R1 (F2, R3)	KEF1/LOR2	LOF2/LOR1	LOF1/LOR1
Microgale jenkinsae	F4/R6 (F2, R7)	LOF1/LOR1	LOF2/LOR1	LOF1/LOR1
Microgale jobihely	F4/R5 (F2, R7, R8)	KEF1/LOR2	LOF2/LOR1	LOF1/LOR1
Microgale longicaudata	F1/R1 (F2, R3)	LOF1/LOR1	LOF2/LOR1	LOF1/LOR1
Microgale majori	F1/R1 (F2, R3)	LOF2/LOR2	LOF2/KER1	LOF1/LOR1
Microgale monticola	F1/R1 (F2, R3)	LOF2/LOR1	LOF2/LOR1	LOF1/LOR1
Microgale nasoloi	F3/R1 (F2, R3)	KEF1/LOR2	LOF2/KER1	LOF1/LOR1
Microgale parvula	F1/R1 (F2, R3)	LOF1/LOR1	LOF2/LOR1	LOF1/LOR1
Microgale principula	F1/R1 (F2, R3)	LOF1/LOR1	LOF2/KER1	LOF1/LOR1
Microgale pusilla	F1/R1 (F2, R3)	KEF1/LOR1 (LOF2, LOR2)	LOF2/LOR1	LOF1/LOR1
Microgale soricoides	F1/R1 (F2, R3)	KEF1/LOR2	LOF2/LOR1	LOF1/LOR1
Microgale taiva	F1/R1 (F2, R3)	LOF1/LOR1	LOF2/LOR1	LOF1/LOR1
Microgale talazaci	F1/R1 (F2, R3)	LOF1/LOR1	LOF2/KER1	LOF1/LOR1
Microgale thomasi	F1/R1 (F2, R3)	LOF2/LOR1 (LOR2)	LOF2/LOR1	LOF1/LOR1
Oryzorictes hova	F1/R1 (F2, R3)	LOF1/LOR1	**AJ893568**	**AY986750**
Oryzorictes tetradactylus	F1/R1 (F2, R2)	LOF1/LOR1	LOF2/LOR1	LOF1/LOR1
Setifer setosus	F1/R1 (F2, R3)	LOF1/LOR1	**AJ893566**	LOF1/LOR1
Tenrec ecaudatus	**AJ251108**	LOF1/LOR1	**AJ893564**	LOF1/LOR1

Supplementary Table 1.2 (continued).

Species	BRCA1	GHR	Rag1
Amblysomus hottentotus	**AF284027**	**JN414680**	LOF1/LOR2
Chrysochloris asiatica	**JN414131**	**JN414681**	LOF1/LOR4 (LOF2, KER3)
Chrysochloris stuhlmanni	F1/R1 (F3, KER1)	F1g/R7	LOF1/LOR4 (LOF2, KER3)
Chrysospalax trevelyani	-	**AF392877**	-
Eremitalpa granti	-	-	-
Micropotamogale lamottei	**AY121759**	**DQ202290**	**JQ073179**
Micropotamogale ruwenzorii	-	-	-
Potamogale velox	-	**DQ202291**	-
Echinops telfairi	**AF284025**	**AF392889**	**JQ073181**
Geogale aurita (1)	F1/R1 (F2, R2)	F1g/R7	LOF1/LOR1 (LOF2, KER3)
Geogale aurita (2)	KEF3/GeauR1	**AM905347**	KEF1/LOR3
Hemicentetes nigriceps	F1/R1 (F2, R2)	F1g/R7	LOF1/LOR4 (LOF2, LOR2)
Hemicentetes semispinosus	F2/R1	**DQ202288**	KEF1/LOR1 (KEF2, LOR4, KER3)
Limnogale mergulus	F1/R1 (F2, R2)	**DQ202289**	**JQ073180**
Microgale brevicaudata	F1/R1 (F3, R2)	**AM905345**	LOF1/LOR1 (LOF2, KER3)
Microgale cowani	F1/R1 (KEF3, R2)	F1g/R7	LOF1/LOR1 (LOF2, LOR2)
Microgale dobsoni	F1/R1 (KEF1, KER1)	F1g/R7	LOF1/LOR1 (LOF2, LOR2)
Microgale drouhardi	F1/R2 (KEF3, F3)	F1g/R7	LOF1/LOR1 (LOF2, LOR2)
Microgale dryas	F1/R1 (F4, R3)	F1g/R7	LOF1/LOR1 (LOR2)
Microgale fotsifotsy	F1/R1 (F3, R2)	F1g/R7	LOF1/LOR1 (LOF2, KER3)
Microgale gracilis	F1/R1 (F2, R2)	F1g/R7	LOF1/LOR1 (LOF2, KER3)
Microgale grandidieri	KEF1/R1 (KEF3, KER1)	F1g/R7	LOF1/LOR1 (LOF2, KER3)
Microgale gymnorhyncha	F1/R1 (F2, R2)	F1g/R7	LOF1/LOR1 (LOF2, KER3)
Microgale jenkinsae	KEF1/R1 (KEF3, F2, R3, KER1)	F1g/R7	KEF1/LOR1 (LOF2, KER3)
Microgale jobihely	KEF1/R1 (F2, F4, KER1, R3)	F1g/R7	LOF1/LOR1 (LOF2, LOR2)
Microgale longicaudata	KEF1/R3 (F4, KER2)	F1g/R7	LOF1/LOR1 (LOF2, LOR2)
Microgale majori	KEF1/R1 (KEF3, KER1)	F1g/R7	LOF1/LOR1 (LOF1, LOR2, LOR4)
Microgale monticola	F1/R1 (F3, KER1)	F1g/R7	LOF1/LOR1 (LOF2, LOR2, KER3)
Microgale nasoloi	F1/R1 (F2, R2)	F1g	LOF1/LOR1 (LOF2, LOR2)
Microgale parvula	F1/R1 (F2, R2)	F1g/R7	LOF1/LOR1 (LOF2, LOR2)
Microgale principula	KEF1/R1 (KEF2, F2, R3)	F1g/R7	LOF1/LOR1 (LOF2, LOR2)
Microgale pusilla	F1/R1 (F2, R2)	F1g/R7	LOF1/LOR1 (LOF2, LOR2)
Microgale soricoides	F1/R1 (F2, R2)	F1g/R7	LOF1/LOR1 (LOF2, LOR2)
Microgale taiva	F1/R1 (F3, R2)	F1g/R7	LOF1/LOR1 (LOF2, LOR2, KER3)
Microgale talazaci	F4/R1 (F2, R3)	AF392885	LOF1/LOR1 (LOF2, LOR2)
Microgale thomasi	F1/R1 (F2, R2)	F1g/R7	LOF1/LOR1 (LOF2, LOR2, KEF2, KER3)
Oryzorictes hova	F1/R1 (F2, R2)	**AF392886**	LOF1, LOR4 (LOF2, LOR2)
Oryzorictes tetradactylus	F1/R1 (F2, R2)	F1g/R7	LOF1/LOR1 (LOF2, LOR4)
Setifer setosus	F1/R1 (F2, R2)	**DQ202292**	LOF1/LOR1 (LOF2, LOR2)
Tenrec ecaudatus	**AF284026**	**AF392890**	LOF1/LOR1 (KEF2, LOR2)

Supplementary Table 1.2 (continued).

Species	vWF
Amblysomus hottentotus	**U97534**
Chrysochloris asiatica	A/B2 (F1, L3)
Chrysochloris stuhlmanni	A/B2 (F1, L3)
Chrysospalax trevelyani	-
Eremitalpa granti	-
Micropotamogale lamottei	**AF390538**
Micropotamogale ruwenzorii	-
Potamogale velox	-
Echinops telfairi	**AF076478**
Geogale aurita (1)	A/B2 (F1, F2, R2, L3)
Geogale aurita (2)	**AM905352**
Hemicentetes nigriceps	A/B2 (F1, L3)
Hemicentetes semispinosus	**AJ891093**
Limnogale mergulus	**AJ891096**
Microgale brevicaudata	A/B2 (F1, L3, 10R1)
Microgale cowani	A/B2 (F1, L3)
Microgale dobsoni	A/B2 (F1, L3)
Microgale drouhardi	A/B2 (F1, L3)
Microgale dryas	10F1/10R1 (F1, F2, L3)
Microgale fotsifotsy	A/B2 (F1, L3)
Microgale gracilis	A/B2 (F1, L3)
Microgale grandidieri	KHF1/KHR2 (KHF2)
Microgale gymnorhyncha	A/B2 (F1, L3)
Microgale jenkinsae	KHF1/KHR2 (KHF2, KHR1)
Microgale jobihely	KHF1/KHR2 (KHF2, KHR1)
Microgale longicaudata	A/B2 (F1, L3)
Microgale majori	KHF1/KHR2 (F1, KHF2, KHR1, L3)
Microgale monticola	A/B2 (F1, L3)
Microgale nasoloi	A/B2 (F1, L3)
Microgale parvula	A/B2 (F1, L3)
Microgale principula	A/B2 (F1, F2, NF2, 10R1, L3)
Microgale pusilla	A/B2 (F1, L3)
Microgale soricoides	A/B2 (F1, L3)
Microgale taiva	A/B2 (F1, L3)
Microgale talazaci	A/B2 (F1, L3)
Microgale thomasi	A/B2 (F1, L3)
Oryzorictes hova	10R1/B2 (F1, L3)
Oryzorictes tetradactylus	A/B2 (F1, L3)
Setifer setosus	A/B2 (F1, L3)
Tenrec ecaudatus	A/B2 (F1, F2, R2, L3)

Supplementary Table 1.3. Primers used to amplify and sequence each gene (5'-3'). F = forward, R = reverse. All primers were designed by the authors of this study unless otherwise noted.

Gene	Primer	F/R	Sequence (5'-3')	Author
12S+Val	1	F	AAAAAGCTTCAAACTGGGATTAGATACCCCACTAT	Kocher et al. 1989
12S+Val	2	R	TGACTGCAGAGGGTGACGGGCGGTGTGT	Kocher et al. 1989
12S+Val	AF4	F	CTTAAAGGACTTGGCGGT	
12S+Val	AF6	F	AACGTTAGGTCAAGGTGTA	
12S+Val	AR3	R	GCTGAAGATGGCGGTATA	
12S+Val	AR6	R	TGAAATCTTCTGGGTGTA	
12S+Val	C	F	AAAGCAAGGCGCTGAAAAT	Springer and Douzery 1996
12S+Val	G	R	GCGGTACTATCTCTATAGC	Springer and Douzery 1996
12S+Val	LOF1	F	AAGGAGGATTTAGYAGTAA	
12S+Val	LOR1	R	GCTAGTAGTTCTCTGG	
12S+Val	LOR2	R	TGAAGCACCGCCAAGT	
12S+Val	TuF3	F	TAATCGATAAACCCCGATA	Olson et al. 2005
12S+Val	TuR3	R	GACTAGAAAATGTAGCCCAT	Olson et al. 2005
12S+Val	TuR5	R	AAGTGCACYTTCCAGTAC	Olson et al. 2005
A2AB	F1	F	CCCTACTCSGTGCAGGCCAC	
A2AB	F2	F	GCACCCTGCCTCATCATGA	
A2AB	F3	F	CCTTCCTCATCCTCTTCACC	
A2AB	F4	F	CTTCGGCAATGCCCTGGTCA	
A2AB	R1	R	GGGCAGATGGCACCCAGGCT	
A2AB	R2	R	GCTTAGAGTGTCCATTGRCCTC	
A2AB	R3	R	CAAGGGGATGGGGCTGCTTGG	
A2AB	R4	R	CCTTGGCCCTGGGACCTCTGC	
A2AB	R5	R	CAGGCTGTAGCTGAAGAAGA	
A2AB	R6	R	AGAGGACGAAGACACCGATG	
A2AB	R7	R	CGGTTGCTGCGCTTGGCGATCA	
A2AB	R8	R	CAGGTAGATGCGCAGGTAGA	
AQP2	KEF1	F	ATGTGGGAGCTCCGGTCCAT	
AQP2	KEF2	F	ATGTGGGAGCTCCGGTCCAT	
AQP2	KER1	R	GCATTGACAGCCAGGTCCC	
AQP2	KER2	R	GCATTGACAGCCAGGTCCC	
AQP2	LOF1	F	TTTGCAGCATGTGGGAGCTC	
AQP2	LOF2	F	CGGTCCATAGCCTTCTCC	
AQP2	LOR1	R	GAGAGCATTGACAGCCAGGT	
AQP2	LOR2	R	TCGGATGTCSGGTGGCGTGAG	
AR	KER1	R	CCCTGAGGTGGCCGAGTGTAA	
AR	LOF2	F	CCATCTCAGACAGTGCCAAGGA	
AR	LOR1	R	ACTGGAGCTGGGATATGGCA	
BDNF	KEF1	F	GGTTATTTCATACTTCGGTTGCA	
BDNF	KER1	R	CATACACAGGAAGTGTCTATCCT	
BDNF	LOF1	F	GGTTATTTCATACTTCGGTTGCA	
BDNF	LOR1	R	CATACACAGGAAGTGTCTATCCT	

Supplementary Table 1.3 (continued).

Gene	Primer	F/R	Sequence (5'-3')	Author
BRCA1	F1	F	GACTGAATGTAGAAAAGGCTG	
BRCA1	F2	F	TCCTTAACTCGAGCCATGTAAC	
BRCA1	F3	F	GGAAAACGTACCGGAGGAAG	
BRCA1	F4	F	TCTGTAACGAAAGCCAACAG	
BRCA1	GeauR1	R	TTCYCCTCAGCCTGTTC	
BRCA1	KEF1	F	GKAGAAAAGGCTGGATTCTG	
BRCA1	KEF2	F	TGGCTCCTTGCTAAGCCA	
BRCA1	KEF3	F	TGGGAAAACGTATCGGAGGAAG	
BRCA1	KER1	R	AAAGCCAATCAGACGGAGC	
BRCA1	KER2	R	CCCTGAGGACTTYATCAAG	
BRCA1	R1	R	AGCTCTTCACTGCTAGGCCA	
BRCA1	R2	R	GCTTGCTTGATGAAGTCCTCGGG	
BRCA1	R3	R	CCTGCAACCTGAGGACTT	
GHR	F1g	F	GAATTCAACAATGATGACTCTTGG	Asher and Hofrieter 2006
GHR	R7	R	TGTTCAGTTGGTCTGTGCTCAC	Asher and Hofrieter 2006
ND2	1	F	CTAATAAAGCTTTCGGGCCCATAC	Kirchman et al. 2001
ND2	2	R	GCCTTCAAAGCCTTAAGTAGAA	Kirchman et al. 2001
ND2	4	R	TTCCACTTCTGAGTACCAGAAGT	
ND2	2T	F	GACCAAGAGCCTTCAAAGC	
ND2	3TX	F	TAGCMCCATTYCACTTCTGA	
ND2	3TY	F	ACTAGGCATAGCCCCATTCCACTT	
ND2	BrF1	F	ATTGGTGGMTGAGGAGGA	
ND2	F1AG	F	CATACCCCGAAAATGTTGGT	A. Gunderson (pers. comm.)
ND2	LOR2	R	TACGDAAAATCATAGCCTACTC	
ND2	LOR3	R	GCRTARTCATCCATCGCCCACAT	
ND2	MilaF	F	GGCTGAGGAGGGCTAAACCA	
ND2	MilaR1	R	AGGCCTCCCTCCCTTCACCGGATTCTTGC	
ND2	MR1	R	GGATGAATYGCAGCAGTA	
ND2	MR2	R	CTAGCCCCCCTTTCAATCATATACC	
ND2	MR3	R	TAACATCCGCTGTTTTATCAAT	
ND2	NN	F	AAATAAGCTATCGGGCCCATACCCCG	
ND2	NN2	F	CTATCAAAGTAACTCTTTTYTCAGAC	
ND2	PF2	F	TCAATYGCCCACATAGGCTGAAT	
ND2	PotR	R	CCWAAATGANTAATRAT	
ND2	PoveF	F	CATCTGGTAAYTGAAACCT	
RAG1	KEF1	F	GTGTCYTCCAGCACGGA	
RAG1	KEF2	F	GAGGCCATGAAGRGCAGC	
RAG1	KER3	R	GAAGGCCTBGAAGCTTC	
RAG1	LOF1	F	AGCCATCACAGGGAGGCAGA	
RAG1	LOF2	F	GGCATCCGCAGGACCTTCAAGT	
RAG1	LOR1	R	GAATTTGGAGGTGTACAGCCA	
RAG1	LOR2	R	CGTAGCGCTCCAGGTTCTCG	
RAG1	LOR3	R	GGTGGCRTCRCACAGGGT	
RAG1	LOR4	R	TTGCCCTCRTAGCGATACTTG	

Supplementary Table 1.3 (continued).

Gene	Primer	F/R	Sequence (5'-3')	Author
vWF	10F1	F	CGGCCCGGTGGAGCCAGGA	
vWF	10R1	R	GTTAGGGGCCTGCTCACGGT	
vWF	A	F	CTGTGATGGTGTCAACCTCACCTGTGAAGCCTG	Porter et al. 1996
vWF	B2	R	CTGGTCTACATGGTCACAGGAAACCCTGC	
vWF	F1	F	TGGCCCGGAACTTGGYCCGCTA	
vWF	F2	F	GCCCCAYGCCAGCCTCAA	
vWF	KHF1	F	GGAGCCAGGAGACGTTGC	K.B.P. Hildebrandt (pers. comm.)
vWF	KHF2	F	GTACACGCTGTTCCAGGTCTTCAGC	K.B.P. Hildebrandt (pers. comm.)
vWF	KHR1	R	GGCCRTCTTCTTCTTCAGG	K.B.P. Hildebrandt (pers. comm.)
vWF	KHR2	R	GGTAGTGGAGGGCCAG	K.B.P. Hildebrandt (pers. comm.)
vWF	L3	R	TTGTTCTCAGGGGCCTGCTTCTC	
vWF	NF2	F	GGAGTTCMTGGAGSAGGTGATCC	
vWF	R3	R	AATGAGGCGGAGCTGCTTGA	

Primer References

Asher R.J., Hofreiter M. 2006. Tenrec phylogeny and the noninvasive extraction of nuclear DNA. Syst. Biol. 55:181-194.

Kirchman J.J., Hackett S.J., Goodman S.M., Bates J.M. 2001. Phylogeny and systematics of ground rollers (Brachypteraciidae) of Madagascar. The Auk 118:849-863.

Kocher T.D., Thomas W.K., Meyer A., Edwards S.V., Pääbo S., Villablanca F.X., Wilson A.C. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. P. Natl. Acad. Sci. 86:6196-6200.

Olson L.E., Sargis E.J., Martin R. 2005. Intraordinal phylogenetics of treeshrews (Mammalia: Scandentia) based on evidence from the mitochondrial 12S rRNA gene. Mol. Phylogenet. Evol. 35:656-673.

Porter C.A., Goodman M., Stanhope M.J. 1996. Evidence on mammalian phylogeny from sequences of exon 28 of the von Willebrand factor gene. Mol. Phylogenet. Evol. 5:89-101.

Springer M.S., Douzery E. 1996. Secondary structure and patterns of evolution among mammalian mitochondrial 12S rRNA molecules. J. Mol. Evol. 43:357-373.

Supplementary Table 1.4. GenBank sequences for afrotherian taxa added to the 10-gene dataset for divergence dating. Dashes indicate missing data.

		GenBank accession numbers, partitioned by gene									
Tree Label	**Species**	**12S**	**ND2**	**A2AB**	**AQP2**	**AR**	**BDNF**	**BRCA**	**GHR**	**RAG1**	**vWF**
Aardvark	*Orycteropus afer* (Pallas, 1766)	NC_002078	NC_002078	JQ965670	Y10632	AJ893563	AY011456	AF284030	AF392892	AY011878	U31617
Armadillo	*Dasypus novemcinctus* Linnaeus, 1758	NC_001821	NC_001821	-	Y10637	-	XM_004454268	AF484222	-	XM_004483173	AJ278158
Elephant	*Loxodonta africana* (Blumenbach, 1797) or *Elephas maximus* Linnaeus, 1758	NC_000934	NC_000934	JN413853	Y10629	AJ893560	JQ073061	AF284022	AF332012	JN633591	U31615
Hyrax	*Procavia capensis* (Pallas, 1766) or *Heterohyrax brucei* (Gray, 1868)	AB096865	AB096865	JN413842	Y10631	AJ893561	JQ073062	AF284023	AF392896	JN414890	U31619
Macroscelidinean	*Macroscelides proboscideus* (Shaw, 1800) or *Elephantulus rufescens* (Peters, 1878)	NC_004026	NC_004026	JN413834	Y10630	AM905337	AY011454	AF284028	AF332014	AY249884	EU136137
Rhynchocyoninean	*Rhynchocyon petersi* Bocage, 1880 or *Rhynchocyon udzungwensis* Rovero & Rathbun, 2008	EU136153	KF202208	JN413888	-	-	JN633360	JN414132	JN414682	JN414889	JN415025
Sirenian	*Dugong dugon* (Müller, 1776) or *Trichecus manatus* Linnaeus, 1758	NC_003314	NC_003314	-	Y15949	AJ893559	JN633362	AF284019	JN414684	JN633590	U31608

Taxonomic References

Blumenbach J.F. 1797. Handbuch der Naturgeschichte. Göttingen, Germany: Dieterich. p. 125.

Bocage J.V.B. 1880. Notice sur une nouvelle espèce du genre Rhynchocyon, Peters. J. Sci. Math. Phys. Nat. Lisboa 1:159.

Gray J.E. 1868. Revision of the species of Hyrax, founded on the specimens in the British Museum. J. Nat. Hist. 1:44.

Linnaeus, C. 1758. Systema naturæ per regna tria naturæ, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. Tomus I. Editio decima, reformata. Stockholm, Sweden: Laurentius Salvius. pp. 33,34,51.

Müller P.L.S. 1776. Linne's Vollstand. Natursyst. Suppl. Nürnberg, Germany: Raspe. p. 21.

Pallas P.S. 1766. Miscellanea Zoologica. The Hague, Netherlands: Petrum van Cleef. pp. 30,64.

Peters W.C.H. 1878. Über die Hrn. J. M. Hildebrandt während seiner letzten ostafrikanischen Reise gesammelten Säugethiere und Amphibien. Monatsb. K. Preuss. Akad. Wiss. Berlin:198.

Rovero F., Rathbun G. B. 2006. A potentially new giant sengi (elephant-shrew) from the Udzungwa Mountains, Tanzania. Journal of East African Natural History 95:111-115.

Shaw G. 1800. General zoology or systematic natural history: Mammalia. London: White-Friars. p. 536.

Supplementary Table 1.5. Habitat codes used in Ancestral Character State Reconstruction. Species assignments are derived from distribution and habitat data from Goodman et al. (2013). All other afroinsectiphilian taxa (i.e., golden moles and sengis, not shown) were coded as arid-habitat specialists.

Species	Habitat
Echinops telfairi	Arid
Geogale aurita	Arid
Hemicentetes nigriceps	Humid
Hemicentetes semispinosus	Humid
Limnogale mergulus	Humid
Microgale brevicaudata	Arid
Microgale cowani	Humid
Microgale dobsoni	Humid
Microgale drouhardi	Humid
Microgale dryas	Humid
Microgale fotsifotsy	Humid
Microgale gracilis	Humid
Microgale grandidieri	Arid
Microgale gymnorhyncha	Humid
Microgale jenkinsae	Arid
Microgale jobihely	Humid
Microgale longicaudata	Humid
Microgale majori	Humid
Microgale monticola	Humid
Microgale nasoloi	Arid
Microgale parvula	Humid
Microgale principula	Humid
Microgale pusilla	Humid
Microgale soricoides	Humid
Microgale taiva	Humid
Microgale talazaci	Humid
Microgale thomasi	Humid
Oryzorictes hova	Humid
Oryzorictes tetradactylus	Humid
Setifer setosus	Both
Tenrec ecaudatus	Both

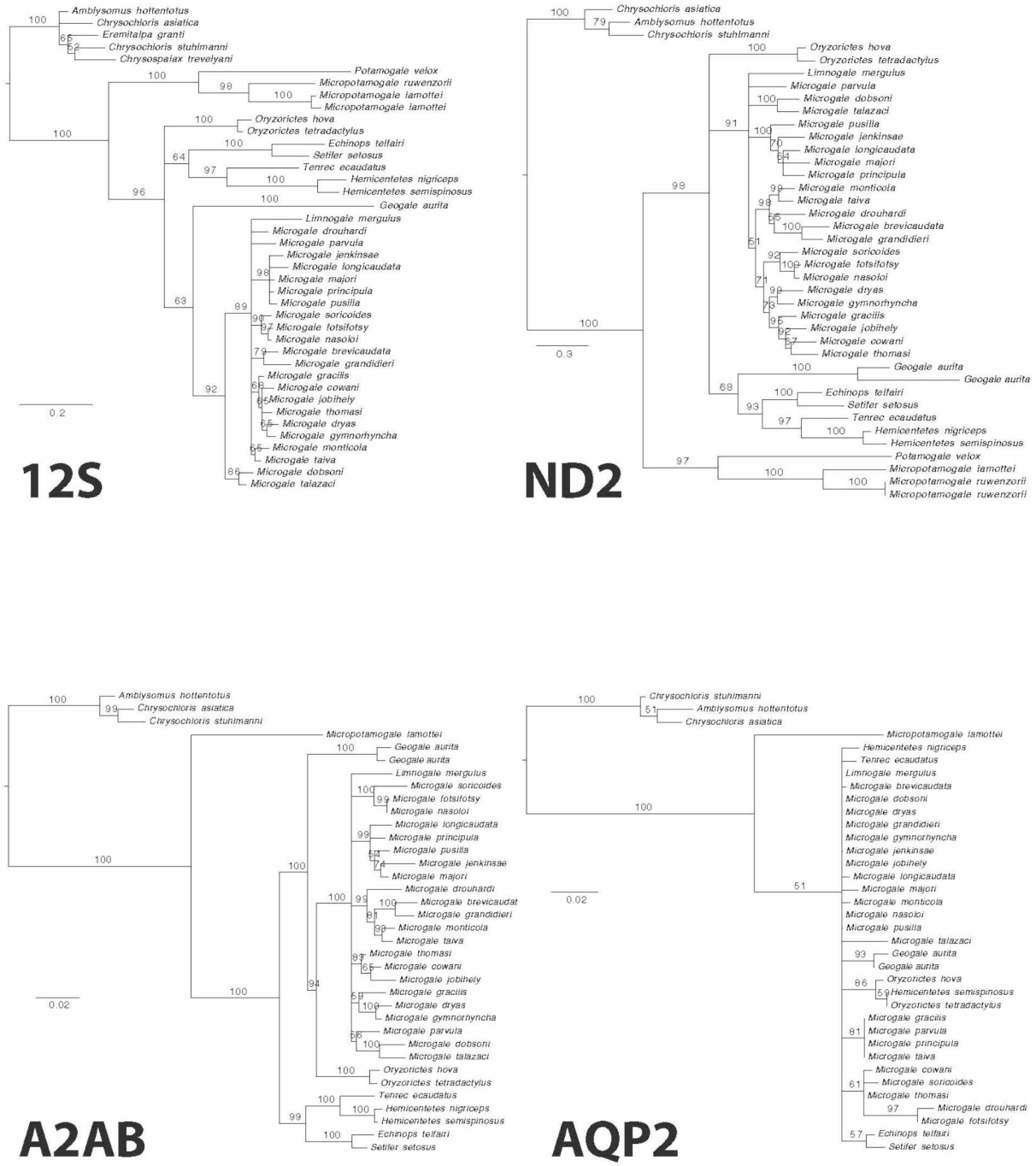
Supplementary Table 1.6. Characteristics of each gene examined in this study. Datasets do not include the non-afrosoricidan sequence data used for divergence dating. Parentheses show alignment statistics with ambiguously aligned characters removed for analyses.

	mtDNA		nuDNA								Combined		
	12S	**ND2**	**A2AB**	**AQP2**	**AR**	**BDNF**	**BRCA1**	**GHR**	**RAG1**	**vWF**	**mtDNA**	**nuDNA**	**All**
Alignment length (unambiguously aligned)	1071 (846)	1044	1131	328	871	648	1026	855	1491	1113	2115 (1890)	7463	9578 (9353)
Percent missing data[a]	0.7% (0.7%)	3.7%	1.9%	4.6%	3.6%	1.3%	5.2%	6.0%	7.7%	1.8%	2.1% (2.4%)	4.2%	3.8% (3.9%)
Parsimony informative sites	424 (257)	649	264	51	180	94	556	245	352	313	1073 (906)	2055	3128 (2961)
Best-fit model of DNA substitution	GTR+ $I+\Gamma$[b]	GTR+ $I+\Gamma$	TVM+I $+\Gamma$	TVM+ $I+\Gamma$	HKY+I $+\Gamma$	GTR+ $I+\Gamma$	GTR+ Γ	K81uf $+\Gamma$	GTR+I $+\Gamma$	TrN+I $+\Gamma$			

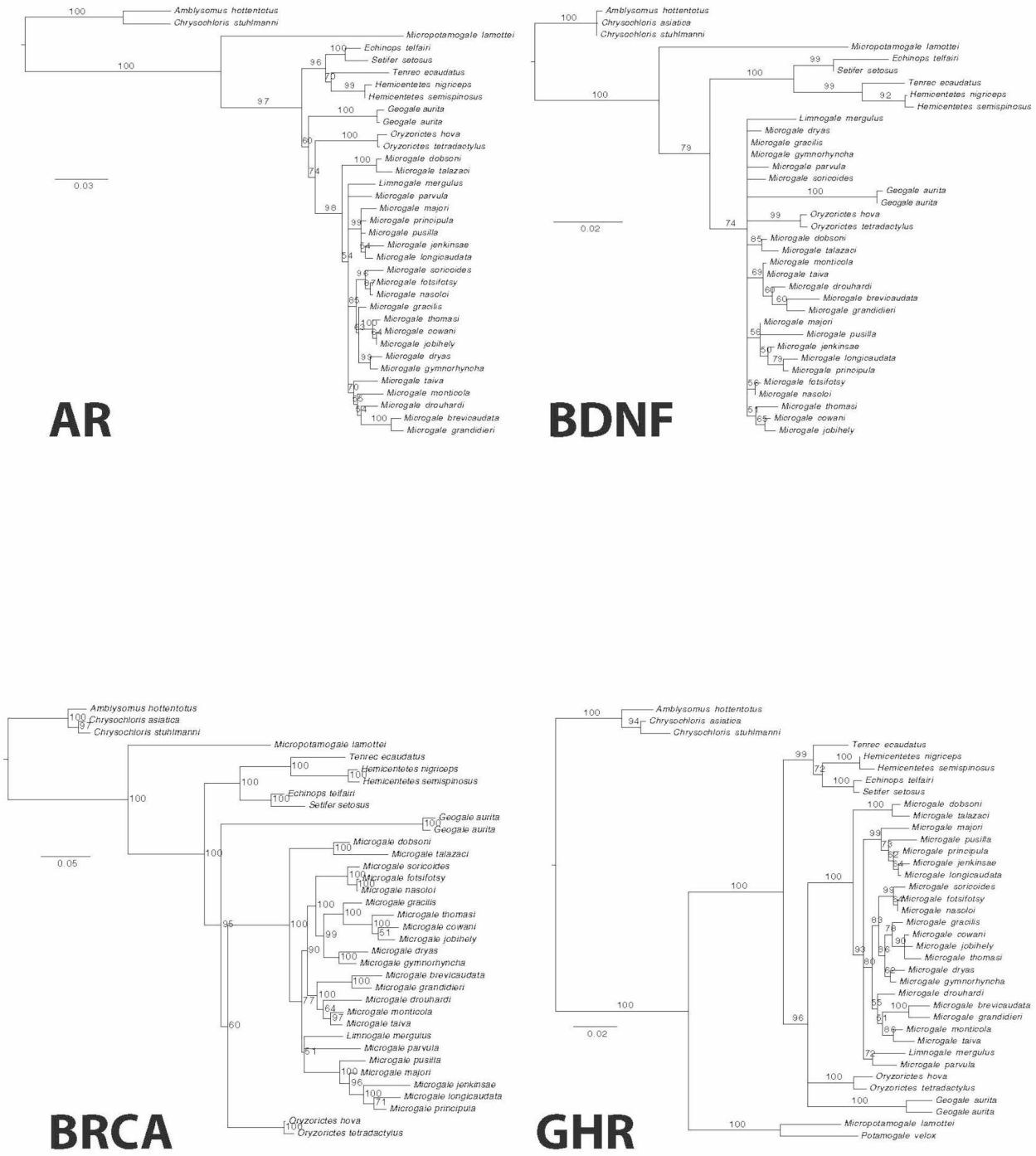
[a]Percentage of the total characters in each data alignment that were unsampled due to incomplete DNA sequences. Interior alignment gaps are not included in this calculation

[b]The pairing or doublet model was applied to stem regions when possible

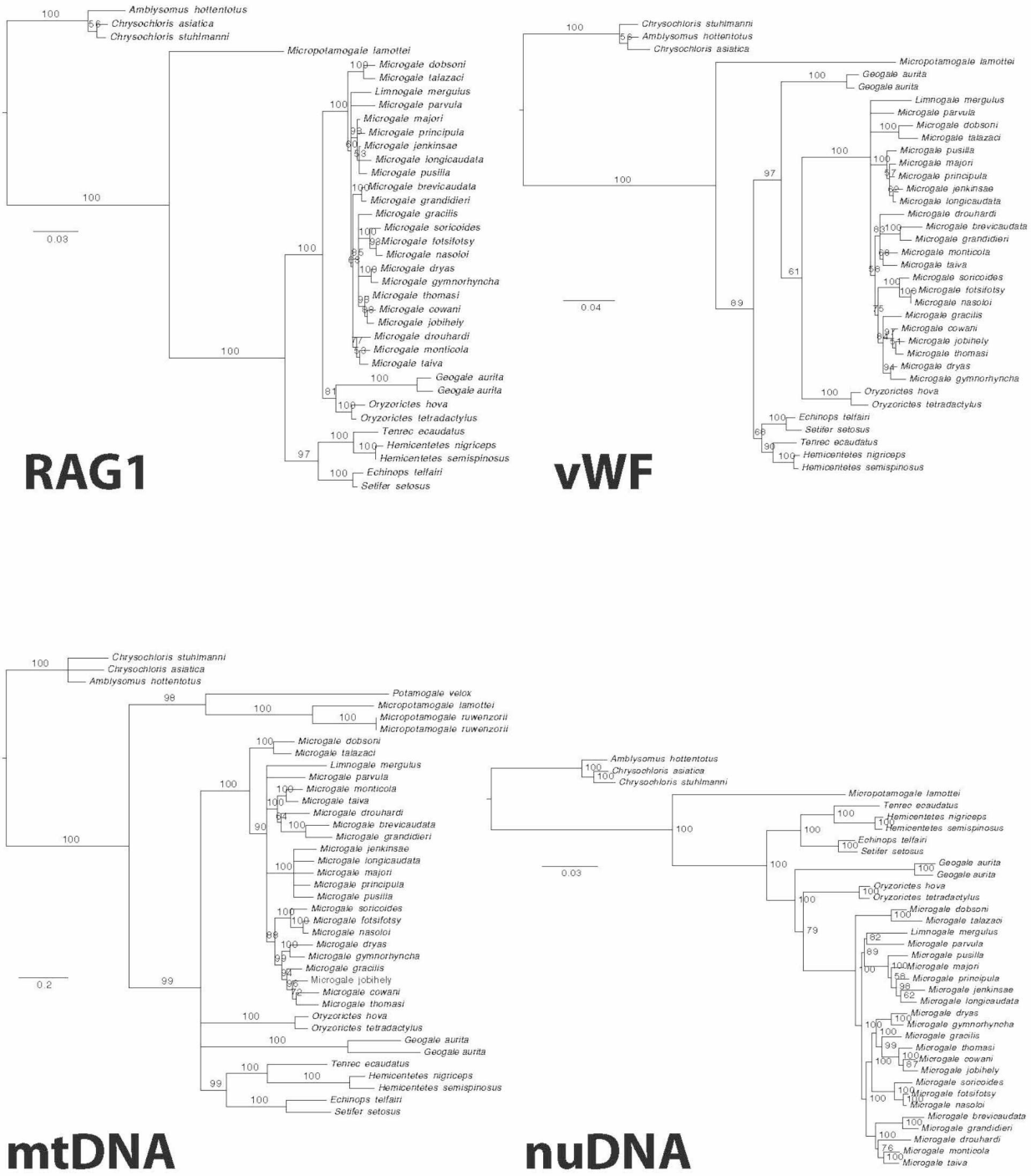
1.11 Supplementary Figures



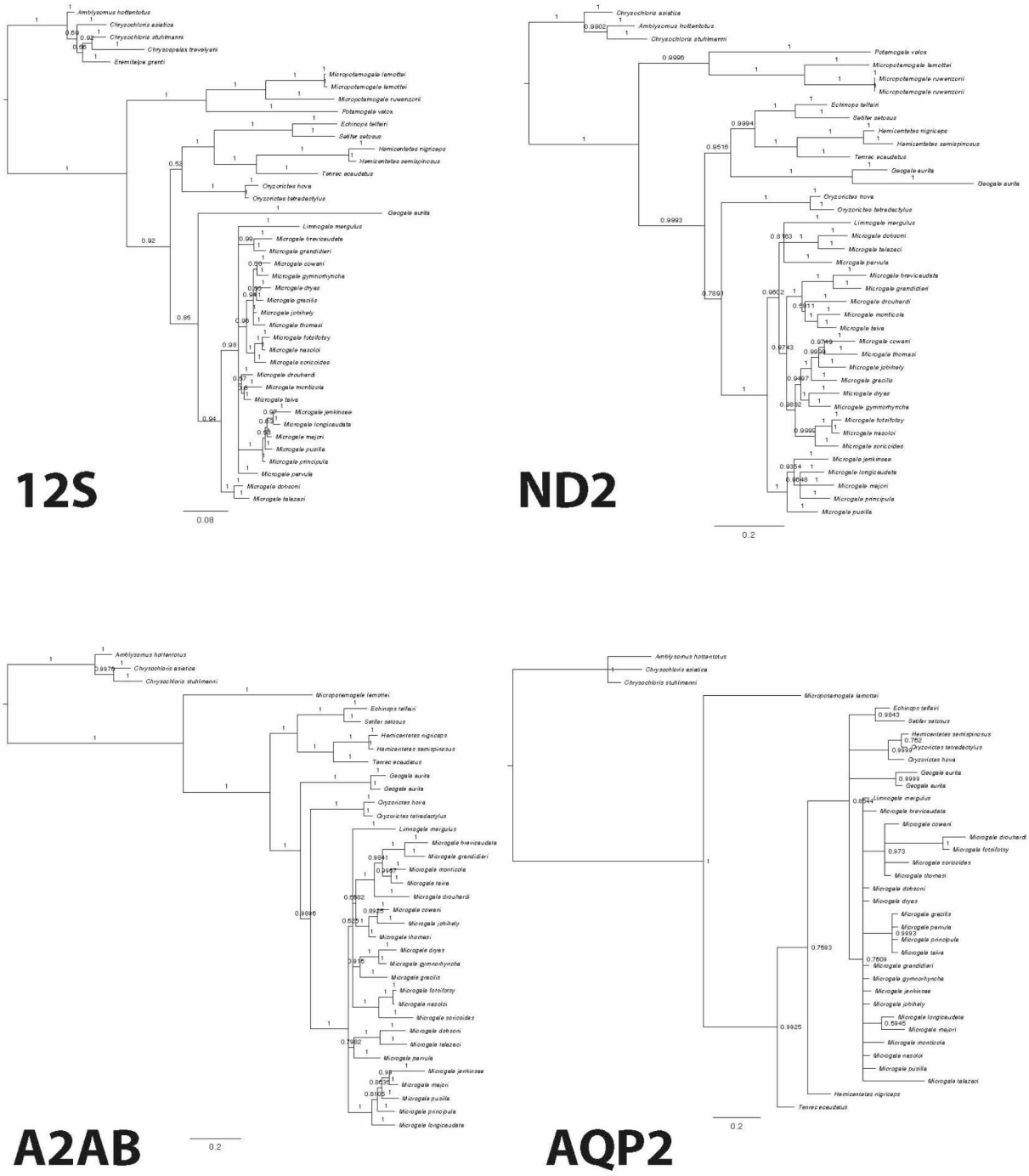
Supplementary Figure 1.1. All individual gene trees produced from ML (GARLI) analyses. Phylogenies from the concatenated mitochondrial (mtDNA) and concatenated nuclear (nuDNA) gene analyses are also provided. Bootstrap support values from 1000 replicates are adjacent to each node.



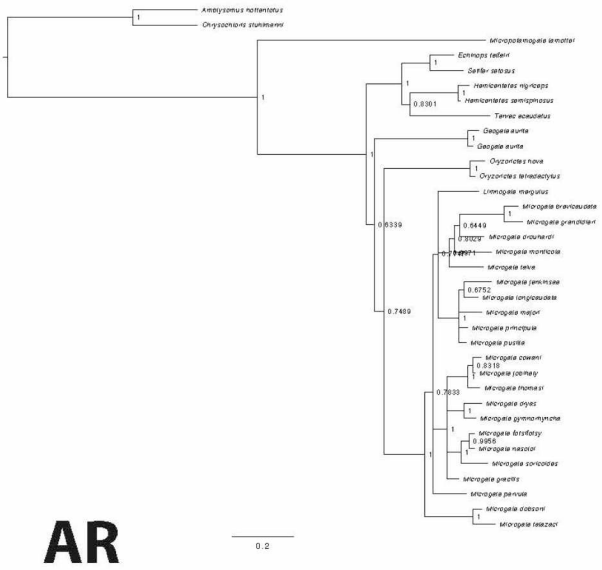
Supplementary Figure 1.1 (continued).

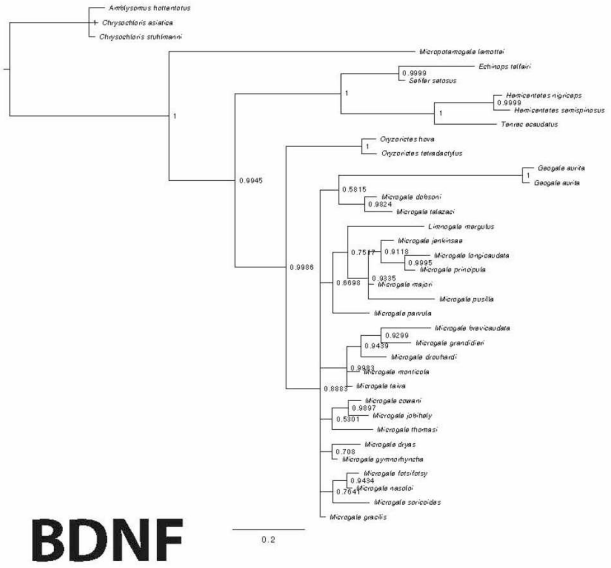


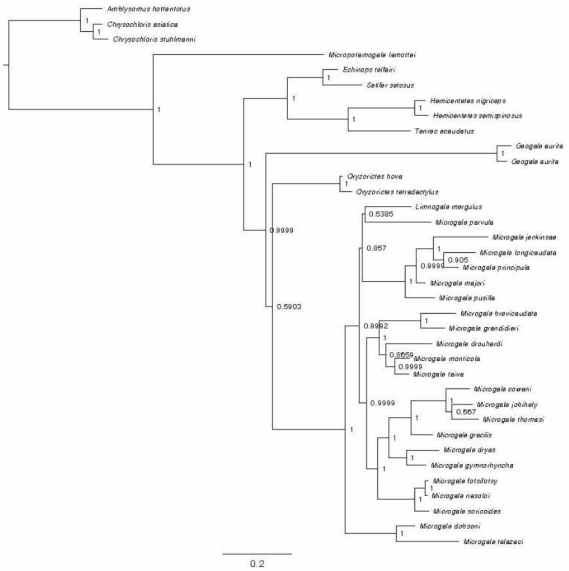
Supplementary Figure 1.1 (continued).

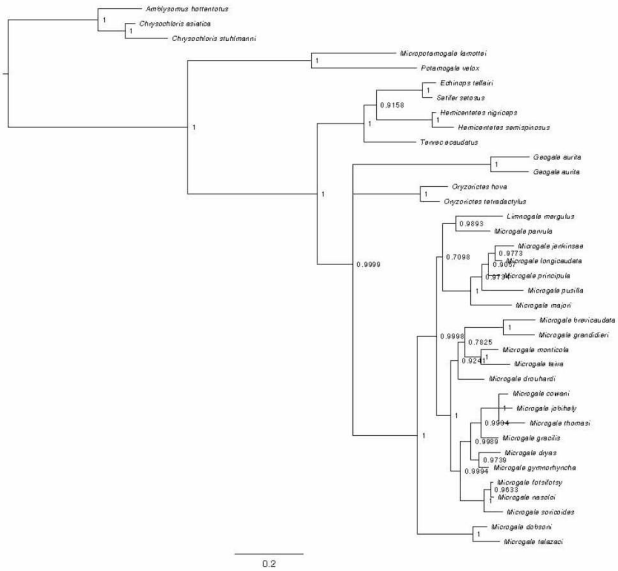


Supplementary Figure 1.2. All individual gene trees produced from Bayesian analyses (MrBayes). Phylogenies from the concatenated mitochondrial (mtDNA) and concatenated nuclear (nuDNA) gene analyses are also provided. Bayesian posterior probabilities are adjacent to each node.



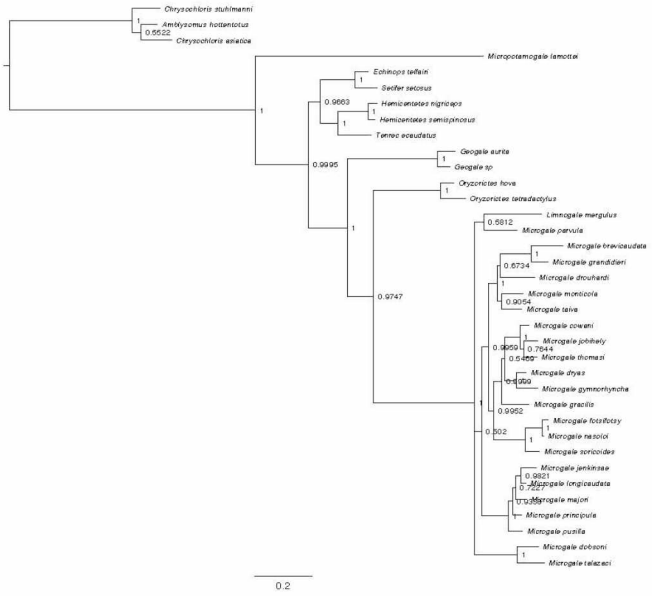




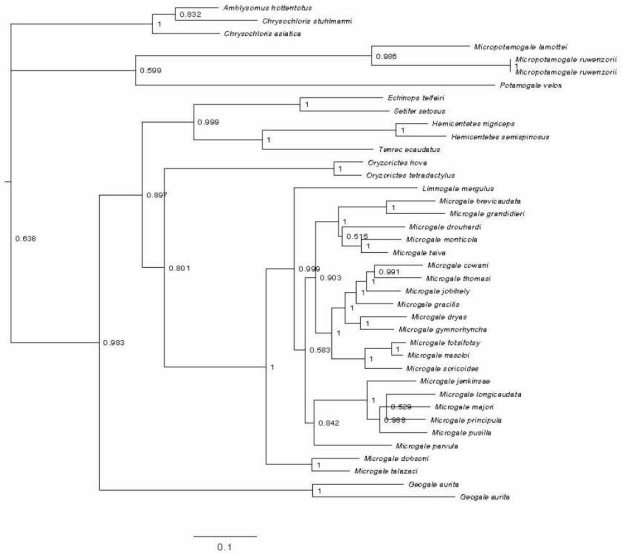


Supplementary Figure 1.2 (continued).

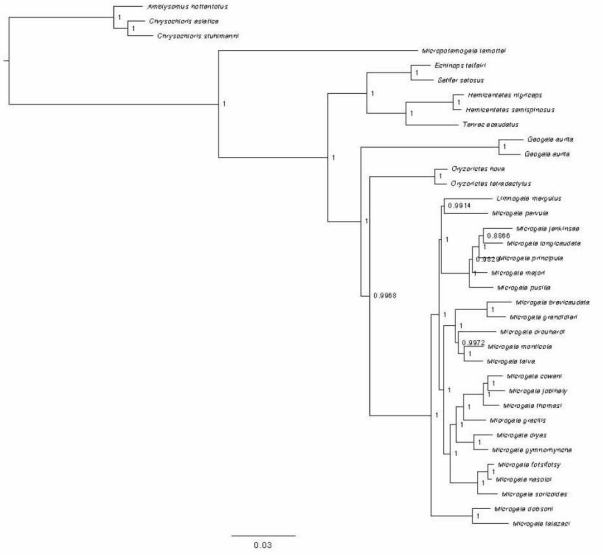
RAG1



vWF

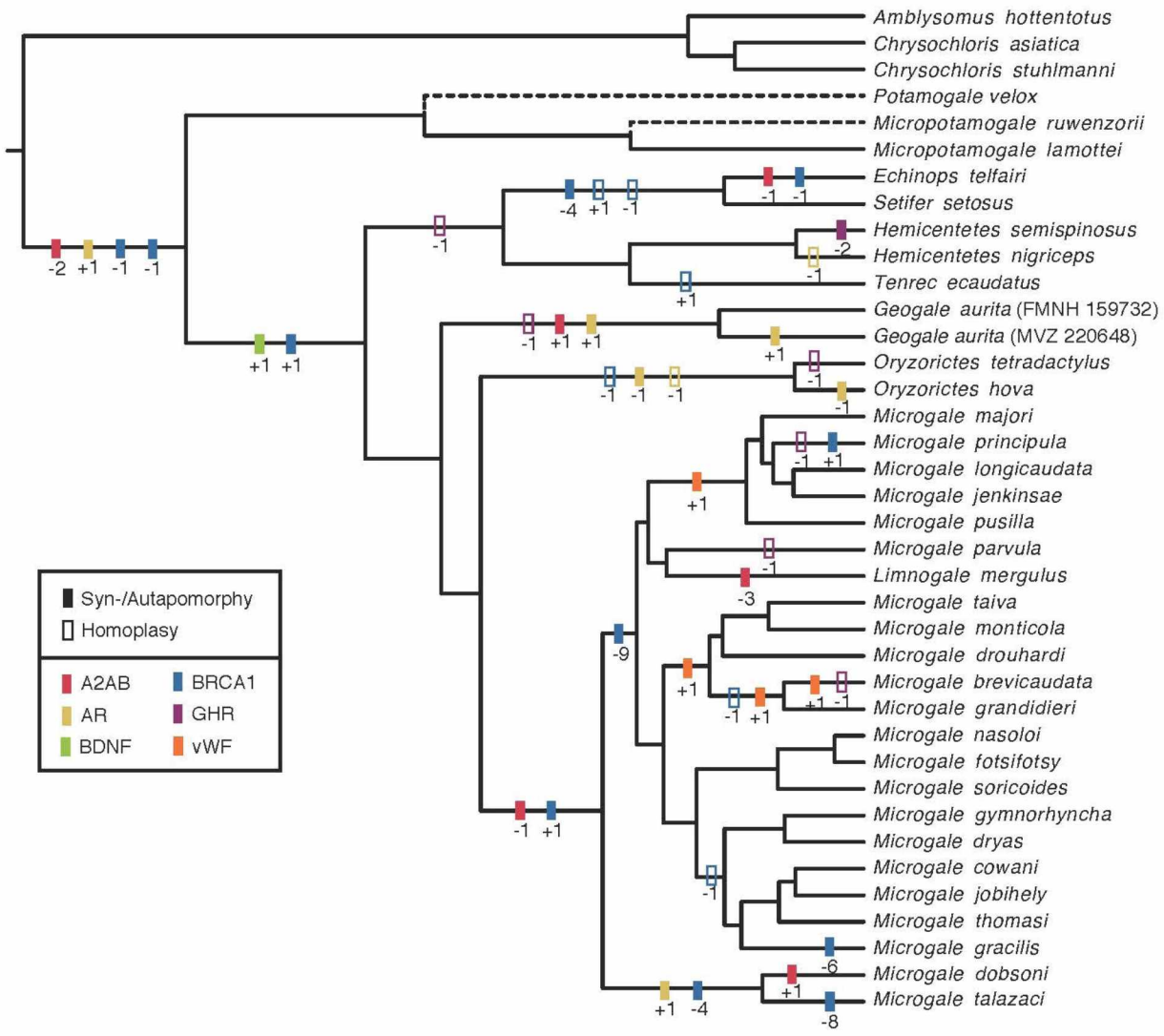


mtDNA

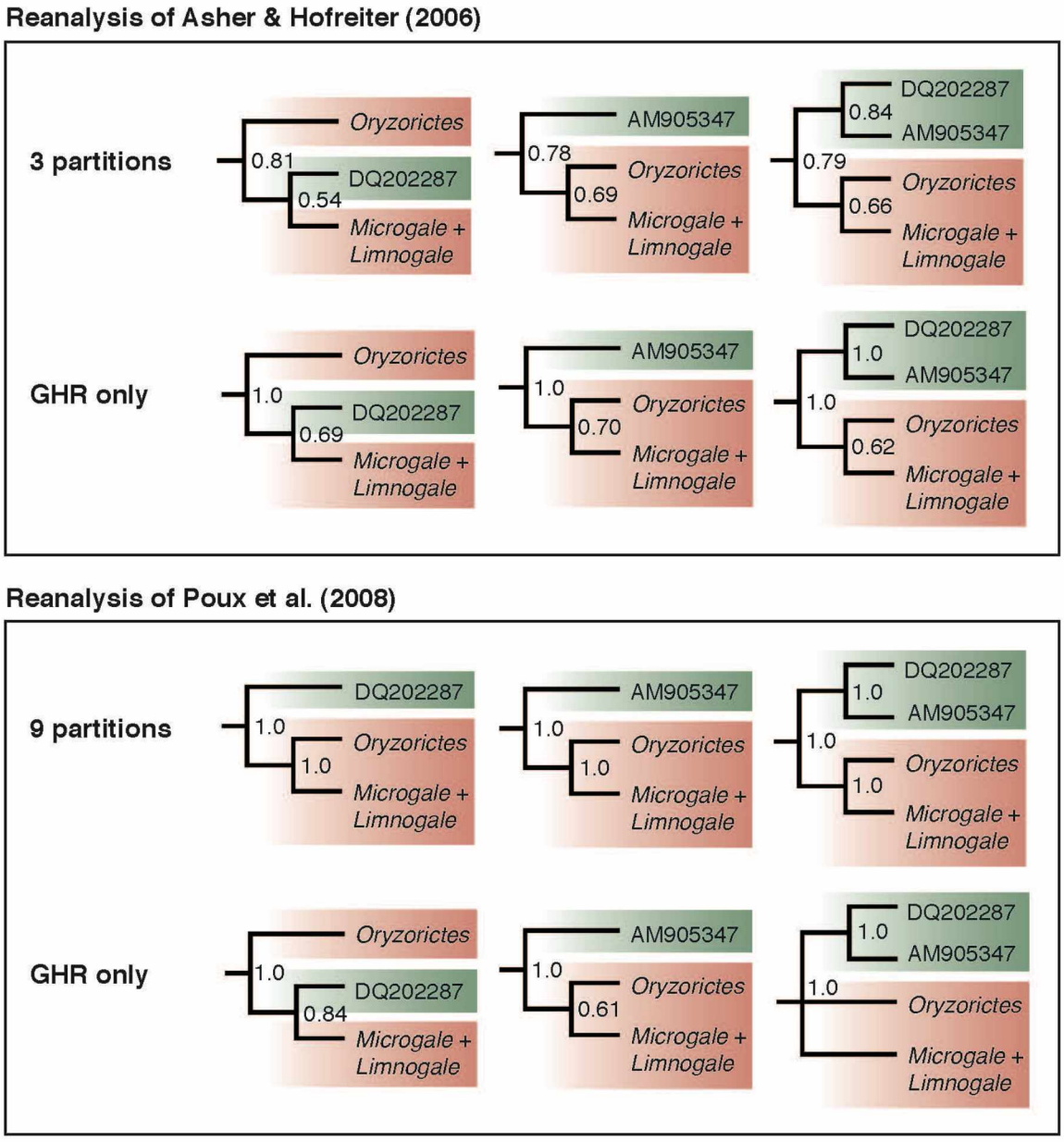


nuDNA

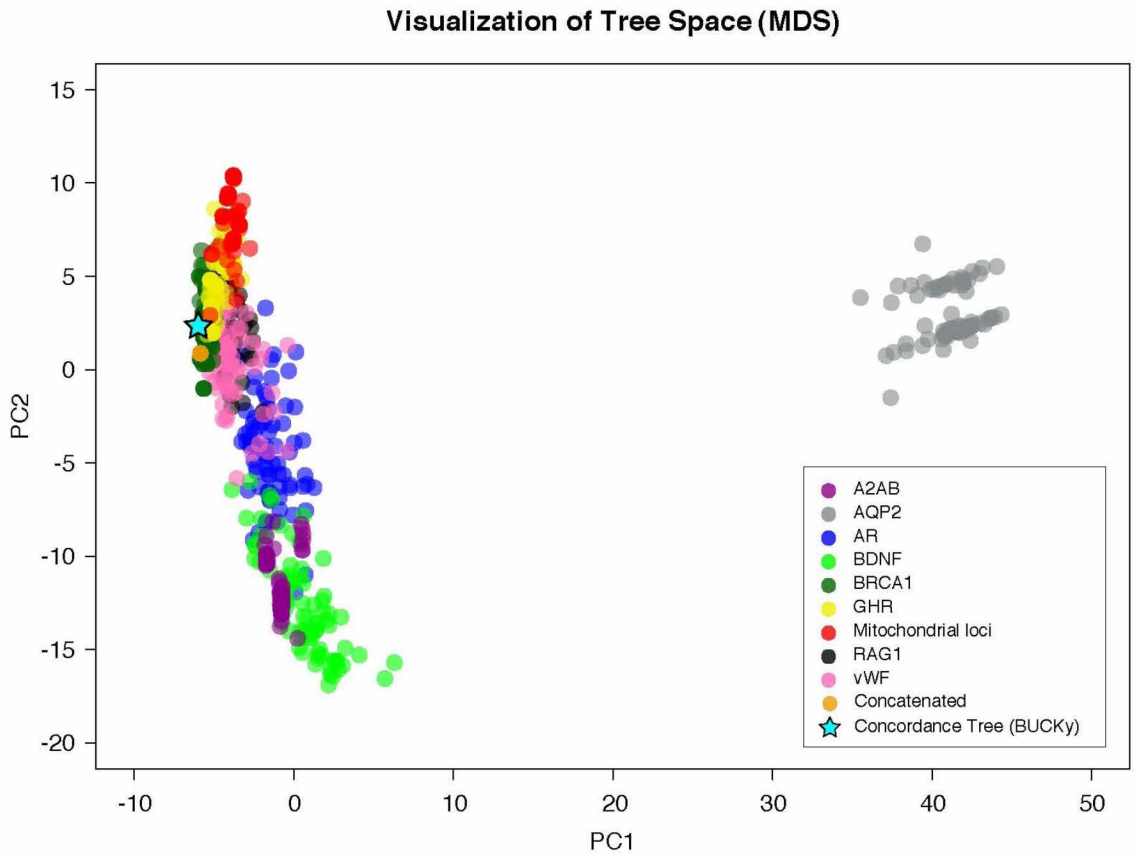
Supplementary Figure 1.2 (continued).



Supplementary Figure 1.3. Insertion-deletion (indel) mutations inferred in the 10-gene concatenated phylogeny. Filled boxes indicate apomorphies, while empty boxes indicate homoplasy. Boxes are color coded by gene, and indel length (bp) and polarity ("+" = insertion, "-" = deletion) are given below boxes. Polarity was determined by comparison to outgroup and sister taxa. Outgroup (golden moles) not shown.



Supplementary Figure 1.4. Majority-rule consensus trees produced by the reanalyses of previously published datasets [i.e., Asher and Hofreiter (2006) and Poux et al. (2008)]. Phylogenies are abbreviated to show the relationships between Geogalinae (green) and Oryzorictinae (red) and the Bayesian posterior probabilities of these relationships. Geogalinae tip labels are the GenBank accession numbers for the *Geogale* GHR sequence published by and analyzed in Asher and Hofreiter (2006) (DQ202287) and Poux et al. (2008) (AM905347). See text for details.



Supplementary Figure 1.5. Multidimensional scaling (principal components analysis) of the Robinson-Foulds distances between phylogenies. We have plotted every tenth tree (after burn-in) from the posterior tree files produced by 10 MrBayes analyses: eight nuclear gene analysis, the full concatenated analysis, and the concatenated mitochondrial analysis. The primary concordance tree from BUCKy is also shown.

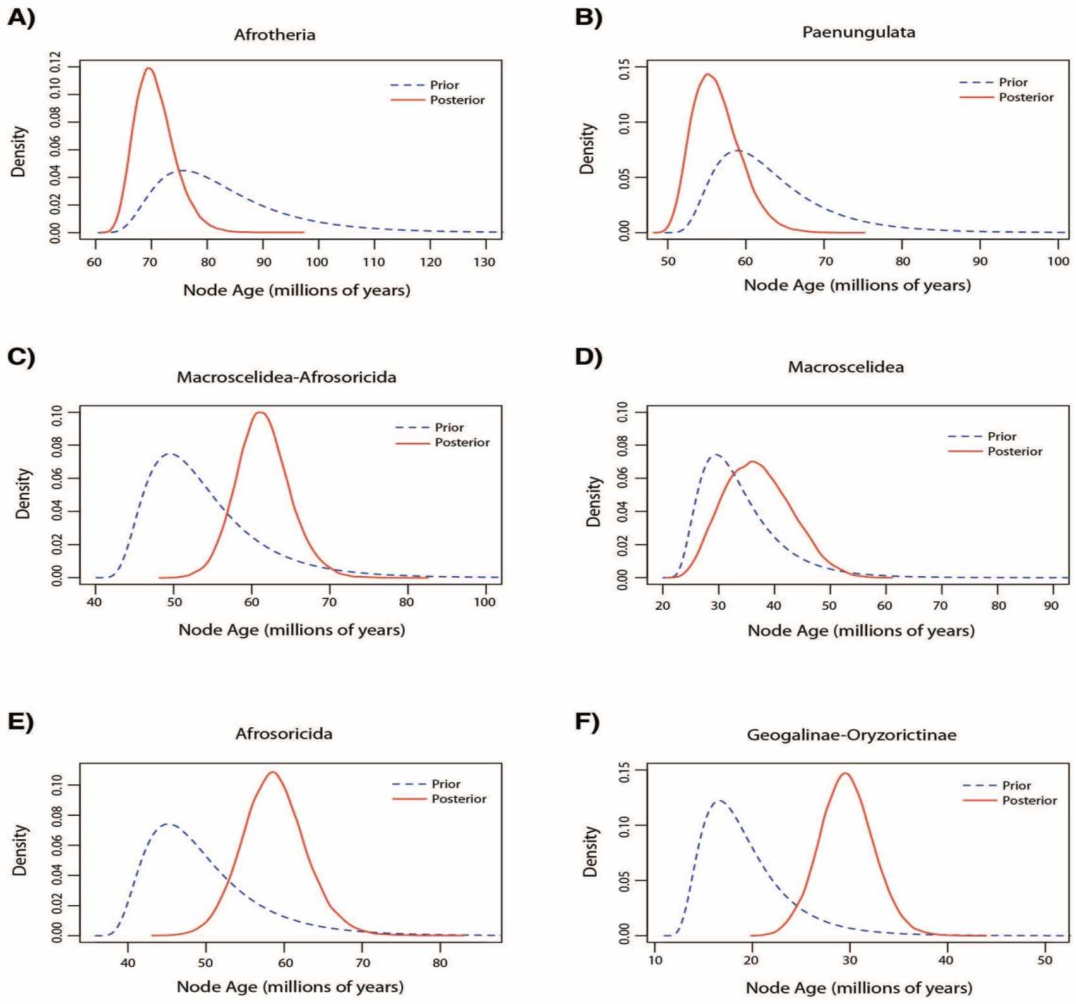
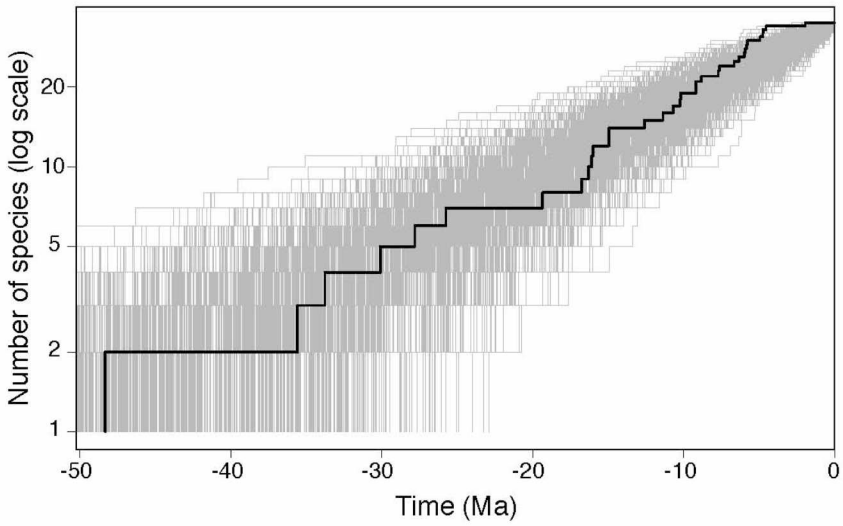
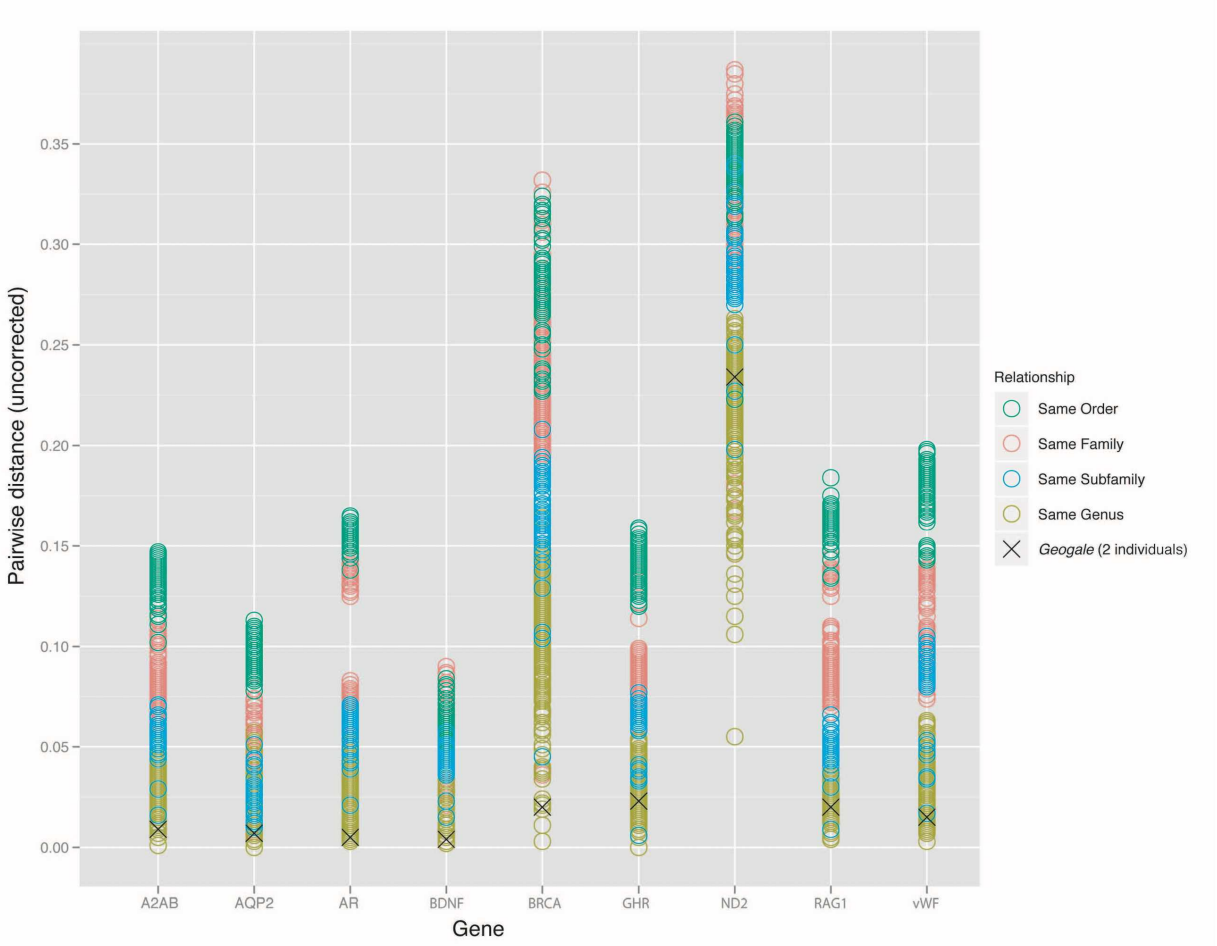


Figure S6. Comparison of prior and posterior distributions on the fossil calibrated nodes used in divergence timing analyses.

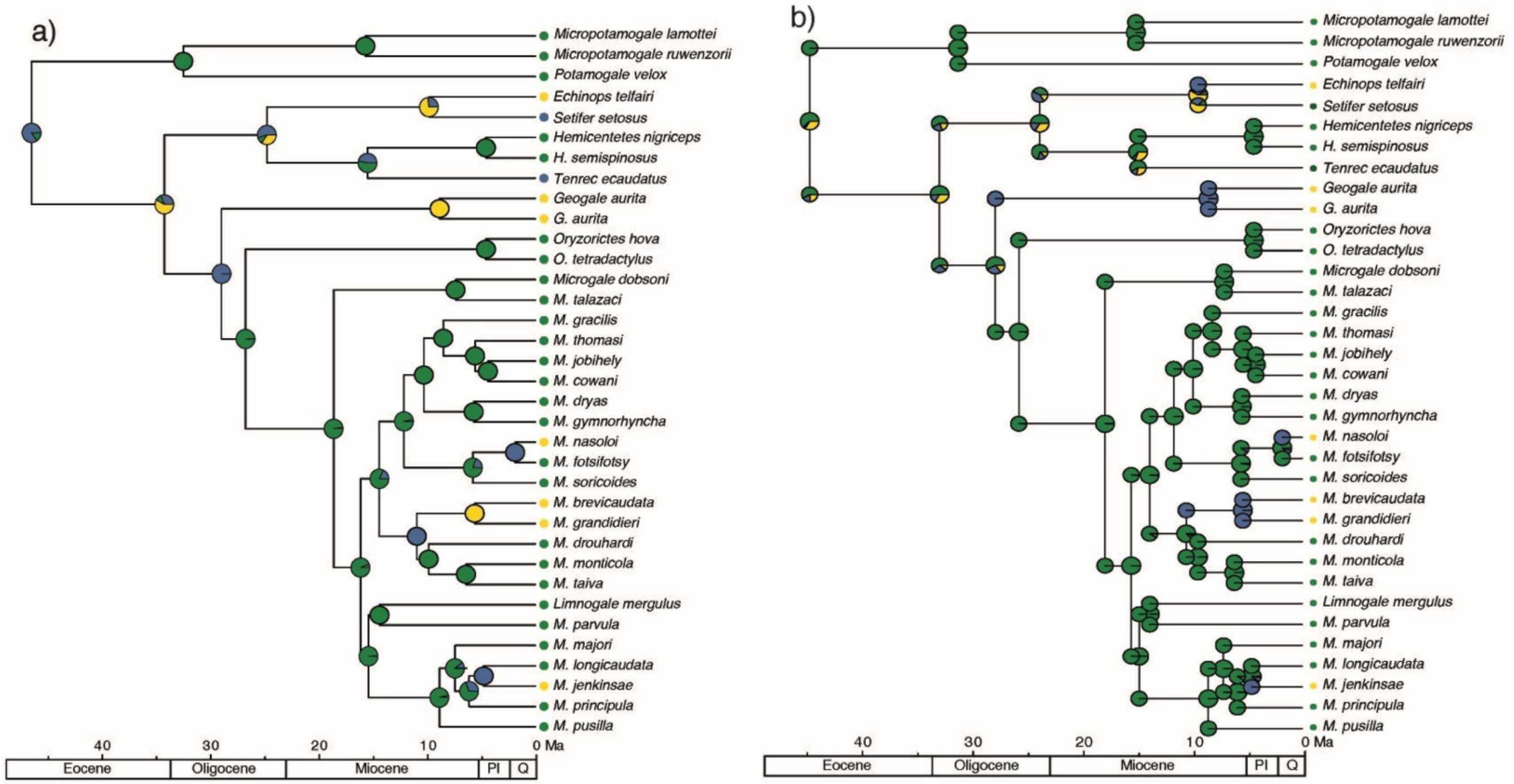
Supplementary Figure 1.6. Comparison of prior and posterior distributions on the fossil-calibrated nodes used in divergence timing analyses.



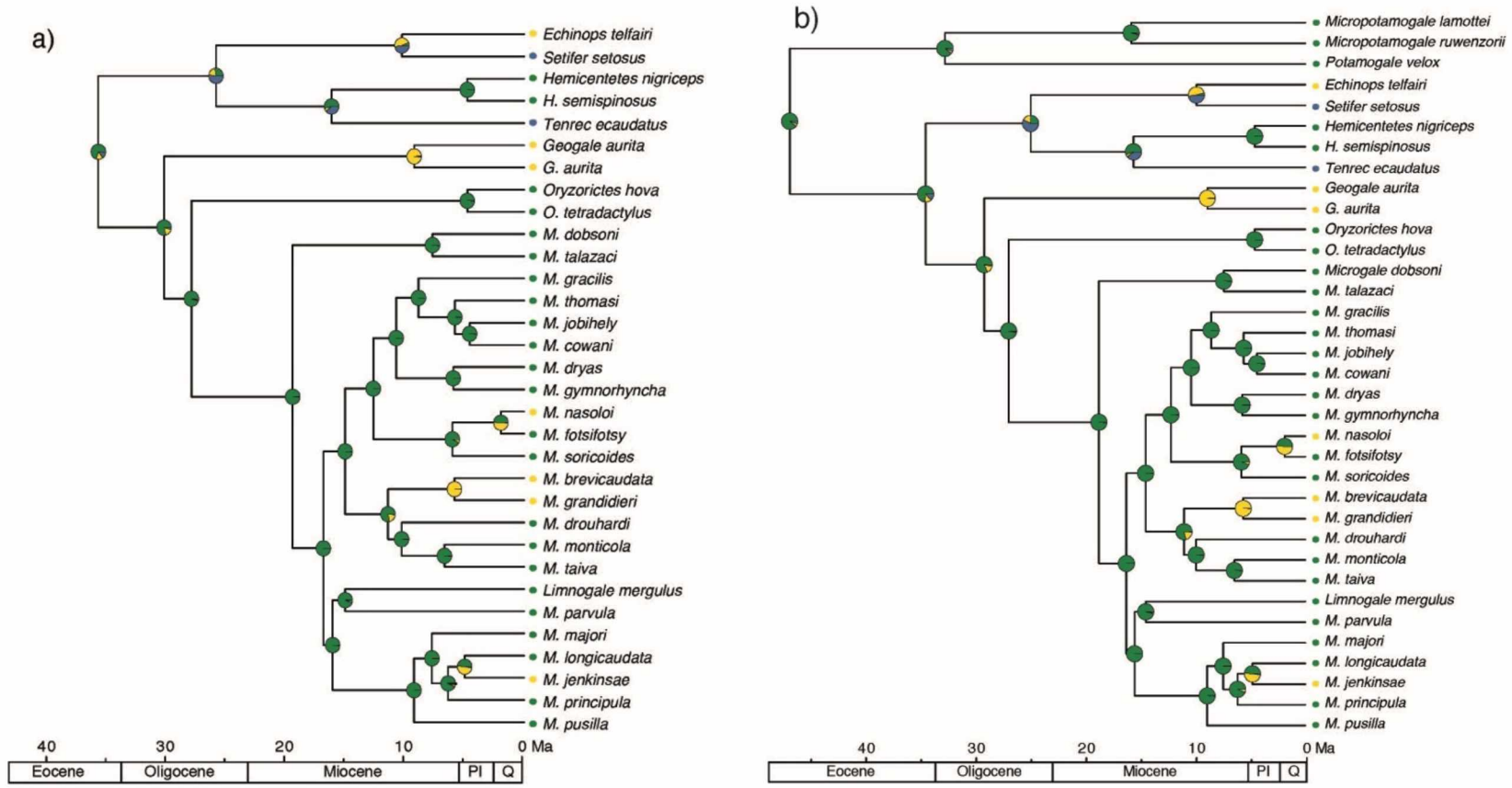
Supplementary Figure 1.7. Observed tenrec lineages through time (black) with lineages through time for 1000 simulated Yule phylogenies (gray).



Supplementary Figure 1.8. Plot of pairwise, uncorrected genetic distances among all species for each gene. Points are color coded to indicate the relationships between the species: species in the same order but different families (i.e., Tenrecidae vs. Chrysochloridae; green), species in the same family but different subfamilies (red), species in the same subfamily but different genera (blue), and species in the same genus (gold). The "X" indicates the pairwise distance between *Geogale aurita* (FMNH 159732) and *G. aurita* (MVZ 220648), which was hypothesized by Poux et al. (2008) to represent a new species. Data from 12S and tRNA-Val are not shown, as we were unable to obtain clean sequences for *G. aurita* (MVZ 220648) for these genes. Relationships reflect the taxonomy prior to revisions in this study (e.g., otter shrews are classified in Tenrecidae in this graph).



Supplementary Figure 1.9. Ancestral character state reconstructions of the habitat associations in Malagasy tenrecs and African otter shrews using a) the GeoSSE model and b) BioGeoBEARs' DEC+j model. Colors correspond to three habitat states: humid (green), dry (yellow), and both/eurytopic (blue). Pie charts show the relative likelihoods of each ancestral habitat. Geological epochs are shown below the timescale. The GeoSSE model reconstructs a single state for ancestral nodes, while BioGeoBEARs reconstructs ancestral nodes as a triplet of parent and daughter states.



Supplementary Figure 1.10. Ancestral character state reconstructions of the habitat associations in a) Malagasy tenrecs and b) Malagasy tenrecs plus African otter shrews using a standard ML model with equal transition rates. Formatting as in Figure S9.

Chapter 2 – Caught in the act: incipient speciation across a latitudinal gradient in a semifossorial mammal from Madagascar, the mole tenrec *Oryzorictes hova* (Tenrecidae)[2]

[2]Everson, K.M., Hildebrandt, K.B.P., Goodman, S.M., & Olson, L.E. (2018). *Molecular Phylogenetics and Evolution* 126: 74-84.

2.1 Abstract

Madagascar is one of the world's foremost biodiversity hotspots, yet a large portion of its flora and fauna remains undescribed and the driving forces of *in situ* diversification are not well understood. Recent studies have identified a widespread, latitudinally structured phylogeographic pattern in Madagascar's humid-forest mammals, amphibians, reptiles, and insects. Several factors may be driving this pattern, namely biogeographic barriers (i.e., rivers or valleys) or past episodes of forest contraction and expansion. In this study, we describe the phylogeographic structure of the small, semifossorial mammal *Oryzorictes hova*, one of Madagascar's two species of mole tenrec, found throughout Madagascar's eastern humid-forest belt, from high-elevation montane forest to low-elevation forests, as well as disturbed habitat such as rice fields. Using one mitochondrial locus, four nuclear loci, and 31 craniomandibular measurements, we identified three distinct populations of *O. hova* associated with the northern, central, and southern regions of the island. We found little evidence of gene flow among these populations, so we treated each population as a potential species. We validated species limits using two Bayesian methods: BP&P, employing only DNA sequence data, and iBPP using both DNA and morphological data, and we assessed whether these methods are susceptible to producing false positive errors. Molecular and morphological data support the recognition of each of the three populations of *O. hova* as distinct species, but formal species descriptions will require additional data from type specimens. This study illustrates the importance of using integrative datasets, multiple methodological approaches, and extensive geographic sampling for species delimitation and adds evidence for a widespread phylogeographic pattern in Madagascar's humid-forest taxa.

2.2 Introduction

Biogeography of Madagascar

Madagascar is renowned for its high levels of endemism and biodiversity. Among native vertebrates, 95% of reptile species, 99% of amphibian species, and 100% of terrestrial mammal species are endemic to the island (Goodman and Benstead 2005, Yoder and Nowak 2006). The humid forest, located along Madagascar's eastern escarpment, coastal plain, and portions of the north and northwest, has long been recognized as a particularly important center of endemism and diversity; an estimated 90% of Madagascar's flora and fauna are strictly forest dwelling (Dufils 2003). Today, Madagascar's humid forests are highly fragmented due to anthropogenic deforestation (Harper et al. 2007, Burns et al. 2016) and possibly due to some natural processes (Quéméré et al. 2012).

Evidence from the fossil record and molecular dating suggest that the majority of Madagascar's plants, invertebrates, and vertebrates are the result of numerous dispersal events from the African continent across the 400-km Mozambique Channel during the Cenozoic (Flynn & Wyss 2003, Poux et al.

2005, Yoder and Nowak 2006). For mammals, a group thought to be fairly adept at overwater dispersal, only four extant native terrestrial lineages are represented (Goodman 2003a): the nesomyine rodents (subfamily Nesomyinae), carnivores (family Eupleridae), lemurs (superfamily Lemuroidea), and tenrecs (family Tenrecidae). Each of these lineages is thought to have been established by single colonization events and subsequent *in situ* diversification (Yoder et al 1996, Yoder et al 2003, Olson and Goodman 2003, Poux et al. 2005). Today there are 27 recognized extant native species of rodents, seven species of carnivores, over 100 species of lemurs, and 31 species of tenrecs on the island of Madagascar, all of which are endemic (Goodman et al. 2013, Mittermeier et al. 2015, Everson et al. 2016, Veron and Goodman 2017, Veron et al. 2017).

Oryzorictes hova

Tenrecs are small- to medium-sized mammals with a striking range of morphological, physiological, and behavioral variation (Olson 2013). A classic example of adaptive radiation, they include semiaquatic, semifossorial, and scansorial ecomorphologies, and body sizes spanning three orders of magnitude (2-2000 grams). As members of the superordinal clade Afrotheria, they are close relatives to golden moles (Chrysochloridae), elephant shrews (Macroscelidea), and aardvarks (Tubulidentata), although their common names reflect ecomorphologies convergent with other mammalian groups (e.g., shrew tenrec, mole tenrec, hedgehog tenrec; Stanhope et al. 1998). Tenrecs [*sensu* Everson et al. 2016, i.e., not including African otter shrews (Potamogalidae)] include eight endemic genera: *Hemicentetes* (streaked tenrecs, 2 extant species), *Setifer (*greater hedgehog tenrec, 1 species), *Echinops* (lesser hedgehog tenrec, 1 species), *Tenrec* (common tenrec, 1 species), *Geogale* (mouse-eared tenrec, 1-2 species; Everson et al. 2016), *Microgale* (shrew tenrecs, 21 species), *Nesogale* (large-bodied shrew tenrecs, 2 species), and *Oryzorictes* (mole or rice tenrecs, 2 species).

The two currently recognized species of mole tenrecs are *Oryzorictes hova* A. Grandidier, 1870 and *O. tetradactylus* Milne-Edwards and A. Grandidier, 1882. Their common name derives from their superficial resemblance to true moles (Talpidae, Talpinae), with small ears and eyes, powerful forelimbs, and elongate claws. The primary characteristic used to distinguish the two species of *Oryzorictes* is the number of digits on the front feet; as its name implies, *O. tetradactylus* has only four, while *O. hova* has five (Goodman 2003b). A third named species, *O. talpoides* G. Grandidier and Petit, 1930, has been synonymized with *O. hova* due to lack of morphological distinctiveness (Goodman et al. 1999). *Oryzorictes hova*, the focal species of this study, is found throughout eastern, northern, and northwestern Madagascar from high-elevation montane forest (nearly 2000 m) to lowland forests (close to sea-level), as well as disturbed habitats such as rice fields (Goodman 2003b, Goodman et al. 2013). It is classified as a species of Least Concern by the International Union for Conservation of Nature (IUCN 2012;

Stephenson et al. 2016a) due to its widespread distribution (Fig. 2.1)—which includes many protected areas—and its occurrence in degraded non-forested habitats. Conversely, *O. tetradactylus* is known from fewer than 10 localities in central Madagascar and is categorized as Data Deficient by the IUCN (Stephenson et al. 2016b).

Phylogeography and species delimitation

Several recent phylogeographic studies of mammals, amphibians, and insects from the humid forests of Madagascar have recovered a consistent pattern of three or more distinct populations along a north-south cline (e.g., Olson et al. 2004, Vieites et al. 2006, Lehtinen et al. 2007, Wirta 2009). These studies have proposed a variety of biogeographic factors that may influence population structure. Wirta (2009) suggested past isolation in tropical forest refugia as an explanation for observed population structure in Malagasy dung beetles. Climatic fluctuations over the last 20 million years would have caused forests to contract and expand, resulting in cycles of isolation and secondary contact for forest-dwelling taxa (Wilmé et al. 2006, Gamisch et al. 2016). Biogeographic barriers such as low valleys or rivers may also underlie population structure. Valleys are known to act as barriers to dispersal for some montane species (e.g., the nesomyine rodents *Voalavo gymnocaudus* and *V. antsahabensis*; Goodman et al. 2005); likewise, rivers serve as biogeographic barriers for some species (Pastorini et al. 2003, Olivieri et al. 2007, Gehring et al. 2012), but are expected to influence montane species to a lesser extent, as rivers are relatively small at high altitudes (Vieites et al. 2006). Continual isolation of populations in allopatry is known to result in speciation (Coyne and Orr 2004), which may or may not involve morphological divergence (e.g., Olson et al. 2004, Lehtinen et al. 2007, Wirta 2009).

Species delimitation, i.e., determining the taxonomic boundaries and numbers of species from empirical data (de Queiroz 2007), is critically important on Madagascar, where forest habitats are rapidly disappearing (Harper et al. 2007) and species-level biodiversity is still vastly underestimated (Vieites et al. 2009, Perl et al. 2014). In recent years, there has been an explosion of new methods that can be used to detect (e.g., O'Meara 2010, Huelsenbeck et al. 2011) or validate (e.g., Yang and Rannala 2010) species limits using genetic data. At the same time, many authors have pushed for renewed integration of multiple lines of evidence, such as molecular, morphological, geographic, and ecological data (e.g., Carstens et al. 2013), and have shown that a so-called "integrative taxonomy" produces more robust species boundaries (e.g., Guillot et al. 2012, Edwards and Knowles 2014, Solís-Lemus et al. 2014). We operate here under the general lineage species concept (GLSC; de Queiroz 2007), which conceptualizes species as separately evolving metapopulation lineages, and we consider genetic monophyly and morphological distinctiveness (both as indicators of reproductive isolation) as different lines of evidence for lineage independence.

As currently recognized, *O. hova* consists of a single, broadly and continuously distributed, relatively common species with no evidence of constituent cryptic lineages. It may therefore represent a 'null' model of phylogeographic structure (or lack thereof) for similarly distributed taxa in Madagascar's humid forests. Some studies have suggested that *O. hova* may be a poor disperser and/or require large, intact blocks of suitable habitat (Goodman and Rakotondravony 2000), but given that this species is documented to occur in degraded forest and different forms of open marshy habitat, it is not considered to be forest dependent. However, limited gene flow during past episodes of forest and marsh habitat contraction may have resulted in inter-population divergence and, possibly, speciation. Using museum specimens from throughout the known range of *O. hova*, we collected DNA sequence data and morphological measurements to characterize phylogeographic structure and identify potential new species. We take a two-step approach to species delimitation, first using unguided methods (phylogenetics, population genetics, and morphological and molecular clustering) to identify and assign individuals to groups or potential species, and then testing the integrity of these groups using species validation methods.

2.3 Materials and Methods

Molecular data collection

We collected DNA sequence data from 43 *Oryzorictes hova* specimens spanning its known geographic range, including 523 base pairs (bp) of the protein-coding mitochondrial gene NADH dehydrogenase subunit 2 (ND2) and 4 nuclear exons: 367 bp from exon 18 of the amyloid beta A4 precursor protein (APP), 617 bp from exon 11 of the breast and ovarian cancer susceptibility 1 gene (BRCA1), 400 bp of recombination activating gene 2 (Rag2), and 410-530 bp of exon 28 of the von Willebrand Factor precursor (vWF; Supplementary Table 2.1, Fig. 2.1). We also sequenced all five markers from one individual of *O. tetradactylus*, the sister species of *O. hova* (Everson et al. 2016), to root our phylogenies. These samples and the associated voucher specimens are housed in the Field Museum of Natural History (FMNH), Chicago, and Mention Zoologie et Biodiversité Animale, Université d'Antananarivo (UADBA, formerly Département de Biologie Animale), Antananarivo.

Genomic DNA was extracted from tissues following the animal tissue protocol from the PureGene DNA isolation kit (Gentra Systems Inc., Minneapolis, Minnesota). Standard PCR amplifications were conducted in 15 μL reactions using primer pairs APP-F1 and APP-R1, BRCA-F1 and BRCA-R2, Met-1 and ND2-LOR2, Rag2-F1 and Rag2-R1, and vWF-F1 and vWF-D (Supplementary Table 2.1). Amplification reactions were electrophoresed and visualized by ethidium bromide staining on a 1.5% sodium borate agarose gel.

We purified 10 μL of the amplification reactions using 0.25 μL exonuclease 1, 0.50 μL shrimp alkaline phosphatase dephosphorylates, and 2.0 μL 10x buffer (USB Corp., Cleveland, OH, USA) at 37 °C for 15 min followed by 80 °C for 15 min. The samples were then cycle sequenced using amplification primers with BigDye Terminator 3.0 and 3.1 (Perkin-Elmer, Boston, MA, USA). All sequences were purified using Sephadex G-50 fine (Amersham Biosciences, Uppsala, Sweden) in a multiscreen filter plate (Millipore Corp., Billerica, MA, USA) and read on an ABI 3130xl Genetic Analyzer (Applied Biosystems, Carlsbad, CA, USA). There were no stop codons or inferred insertion/deletion events in the sequences generated for this study, and sequences were aligned by eye in Sequencher v. 4.7 (Gene Codes, Ann Arbor, MI, USA) with reference to the translated amino sequences of each gene. All generated sequences, along with their specimen voucher numbers, have been deposited to GenBank (accession numbers KX015448 - KX015667; Supplementary Table 2.2).

Morphological data collection

We recorded 31 cranial and mandibular measurements (Supplementary Table 2.3) from 91 vouchered specimens, including the holotypes of *O. hova* and its junior synonym *O. talpoides*. All morphometric data are available upon request, and will be made publicly available upon this chapter's acceptance in a peer-reviewed journal. Most measurements have been employed in previous morphometric studies of other species of tenrecs (Olson et al. 2004, Goodman et al. 2006, Olson et al. 2009). All measurements were recorded by L.E.O. using digital calipers accurate to 0.1 mm connected to a foot pedal and laptop. Only adult specimens, defined by the presence of a fully erupted permanent dentition, were included. Measurements of bilateral elements were taken from the left side whenever possible; when the left-side element was missing or damaged, the right side was measured. See Olson et al. (2004) for additional details.

Phylogenetic analyses

To infer phylogenetic patterns within *O. hova*, we conducted maximum-likelihood (ML) analyses for each gene and for two concatenated datasets (all nuclear loci and all nuclear + mitochondrial loci) in Garli v.0.951 (Zwickl 2006). We used the Akaike Information Criterion (AIC), implemented in ModelTest v.3.7 (Posada and Crandall 1998), to select a best-fit model of nucleotide evolution for each gene. We also used PartitionFinder (Lanfear et al. 2012) to select a partitioning scheme for concatenated analyses. For the PartitionFinder analysis, we used the Greedy algorithm and AIC to identify the optimal number of partitions from 15 possible partitions representing each codon position for the five genes. Run termination and optimization parameters in Garli were set to default values. Nodal support was estimated from 500 bootstrap pseudoreplicates, each of which was allowed to run until $-\ln L$ values converged

(changing less than 0.02) for 2,000 generations. Support values were calculated in PAUP* 4.0 (Swofford 2002) by computing a consensus tree of the 500 bootstrap pseudoreplicates.

We also used the full concatenated dataset to estimate a phylogeny and divergence times using Bayesian Inference (BI) in BEAST v.1.7.5 (Drummond and Rambaut 2007; Heled and Drummond 2010). We used two independent MCMC runs (four chains each) of 10,000,000 iterations each, sampling trees every 1,000 iterations to yield a total of 10,000 trees. Partitions were set according to the optimal partitioning scheme and nucleotide substitution models identified by PartitionFinder. To estimate divergence times, we used the age recovered for the *O. hova – O. tetradactylus* split in Everson et al. (2016) to set a normally distributed prior with a mean of 4.75 MYA and a standard deviation of 1.0. We allowed independent evolution of branch lengths under the uncorrelated relaxed lognormal clock model and the Yule process tree prior. Runs were combined in LogCombiner. The first 20% of these trees were removed as burn-in, and the last 8,000 trees were used to construct a majority-rule consensus tree and assign Bayesian posterior probabilities (BPP) to each branch. Results were summarized with TreeAnnotator v.1.6.1 (Drummond and Rambaut 2007).

Population genetic analyses

We used the Bayesian statistical package Phase v.2.1.1 (Stephens et al. 2001, Stephens and Donnelly 2003) to infer allelic phases of nuclear exons APP, BRCA, Rag2, and vWF from diploid sequence data. Phase requires sequences to be of equal length; therefore, FMNH 166239 and FMNH 178762 were excluded from vWF analysis due to our inability to generate longer sequence fragments. The analysis was allowed to proceed for 10,000 iterations with a burn-in of 1,000 and was repeated a total of three times per exon starting with a different random seed for each run. Unrooted haplotype networks were created in TCS v.1.21 (Clement et al. 2000) using 95% statistical parsimony probabilities for the nuclear exons.

We then used the phased nuclear data to infer the number of populations and the assignment of individuals to those populations using the Bayesian clustering software STRUCTURE v.2.3 (Pritchard et al. 2000). As the STRUCTURE model assumes that input data are independent and unlinked, we employed gene haplotypes rather than individual SNPs. We utilized an admixture model with correlated allele frequencies, and ran all analyses for 500,000 generations after a 100,000-generation burn-in period. We assumed that the number of groups (K) was between 1 and 10, and we performed 10 iterations for each value of K. Results were combined and analyzed in STRUCTURE HARVESTER (Earl and vonHoldt 2012) to determine the optimal value of K using both the peak in the mean probability of the data (Pritchard et al. 2000) and the ΔK method of Evanno et al. (2005). Run files from the optimal K

value were then averaged using CLUMPP (Jakobsson and Rosenberg 2007) and visualized using DISTRUCT (Rosenberg 2003).

As a measure of population differentiation, we estimated fixation indices (F_{ST}) among the genetic clusters identified by STRUCTURE. We also conducted an analysis of molecular variance (AMOVA) to determine the relative proportions of genetic variation within and among populations. Both analyses were calculated in Arlequin v.3.11 (Excoffier and Lischer 2010) using the method of Weir and Cockerham (1984), with levels of significance calculated by a permutational procedure involving 10,000 replications.

If population structure is driven by historical isolation in forest and marsh refugia, as previous authors have hypothesized for different types of organisms (e.g., Wirta et al. 2009, Gamisch et al. 2016), we would expect to see genetic signatures of a recent population bottleneck and subsequent expansion. Thus, we visualized changes in population size through time using the extended Bayesian skyline plot (EBSP), implemented in BEAST v.1.7.5 (Drummond and Rambaut 2007; Heled and Drummond 2010). The EBSP (as compared to the traditional Bayesian skyline plot) permits the analysis of multiple loci with varying rates of coalescence. For each of the populations identified in STRUCTURE, we performed two independent runs of 100 million generations, sampling every 1,000 generations and discarding the first 20 million generations as burn-in. The resulting comma-separated value files (containing columns for time, median theta, and the upper and lower values for the 95% highest posterior density interval of theta) were used to produce skyline plots in R using the package ggplot2 (Wickham 2009).

Morphological analyses

Because museum specimens vary in condition and preparation, not all measurements were possible for all individuals. Therefore, we generated two complete datasets using different methods: (1) the available-case method, whereby individuals or measurements with the greatest amounts of missing data were removed, and (2) imputation of missing data via Multivariate Imputation by Chained Equations using the R package mice (van Buuren and Groothuis-Oudshoorn 2011). The available-case dataset includes 25 craniomandibular measurements and 115 individuals (89 *O. hova* and 26 *O. tetradactylus*), while the imputed dataset includes 31 craniomandibular measurements and 157 individuals (114 *O. hova* and 42 *O. tetradactylus*).

For both species (*O. hova* and *O. tetradactylus*), we confirmed that each craniomandibular measurement was normally distributed using a Shapiro-Wilk test in R (function: shapiro.test, package: stats). We then conducted an analysis of variance (ANOVA) in R (function: aov, package: stats) to test for differences between sexes, species, and clades for each craniomandibular measurement. We also performed principal components analyses (PCA) to evaluate patterns of variation without *a priori*

assignment of individuals to groups. These analyses were performed on both the available-case and imputed datasets using log-transformed (base 10) measurement data.

Species delimitation

Our phylogenetic and population genetic analyses identified three distinct clades (see Sections 3.1 and 3.2), thus we used a variety of methods to evaluate these clades as potential species. First, to test for significant morphological differences between clades, we conducted an ANOVA using the same procedure described in Section 2.4 but including only individuals that had been assigned to a population using DNA sequence data. We then tested whether morphology alone could correctly predict clade membership using a canonical variates analysis (CVA) in which individuals that had already been assigned to clades using molecular data were used as a training dataset. Analyses were performed on both the available-case and imputed datasets on log-transformed (base 10) measurement data using the R package HiDimDA (Silva 2011), which corrects for large numbers of variables using shrunken covariance matrices.

We then used our genetic dataset and the species delimitation software BP&P (Yang and Rannala 2010) to validate species limits. BP&P uses a reversible-jump MCMC algorithm to assign probabilities to speciation events, and the latest version of BP&P (v3.0; Yang and Rannala 2014) simultaneously estimates the species tree. The prior distributions on ancestral population size (Θ) and root age (τ) can influence posterior probabilities; thus, we performed three analyses using different prior combinations for Θ and τ, following the recommendation of Leaché and Fujita (2010). Both priors use a gamma $G(\alpha, \beta)$ distribution: (a) $\Theta = G(1, 10)$ and $\tau o = G(1, 10)$, (b) $\Theta = G(2, 2000)$ and $\tau o = G(2, 2000)$, and (c) $\Theta = G(1, 10)$ and $\tau o = G(2, 2000)$. Large values of Θ and small values of τ favor models with fewer species (Yang and Rannala 2010). Species delimitation analyses (Algorithm 1) were run for 500,000 generations with sampling every 50 generations after a burn-in period of 50,000 generations, resulting in 9,000 saved trees. Convergence was determined by comparing results from two separate independent runs, which were then combined.

Because species delimitation can be improved by considering both molecular and morphological data together (Edwards and Knowles 2014, Solís-Lemus et al. 2014), we used the program iBPP, which uses BP&P's Bayesian framework but can accept independent, continuous measurement data in addition to molecular data (Solís-Lemus et al. 2014). Briefly, this method assumes that each measurement is normally distributed, with species means governed by a Brownian-motion process along the species tree. To meet the assumption of independence in the iBPP model, we did not include all measurements (which may be collinear) but instead used values from the first two canonical variates from the imputed dataset. We employed only those individuals that had already been assigned to clades with molecular data and we

followed the same BP&P parameters described in the previous paragraph. Analyses were conducted using craniomandibular measurement data alone and using both measurement and sequence data together.

Finally, we tested the sensitivity of both BP&P and iBPP to false positives by repeating all analyses using randomized individual-to-species map files, which simulated a panmictic population. We expected that both programs would recover zero or very low posterior probabilities of speciation between hypothesized lineages when individual assignments were randomized. We randomized tip labels using the "sample" function in R and conducted 10 replicates for each analysis.

2.4 Results

Phylogenetic analyses

PartitionFinder identified the best partitioning scheme for concatenated analyses as a single partition containing all loci with the TrN+I model of nucleotide evolution. ML and BI analyses of the concatenated mitochondrial + nuclear dataset (Fig. 2.2) recovered three well-supported clades within *Oryzorictes hova* that correspond to northern, central, and southern regions of Madagascar; these clades will hereafter be referred to as the North, Central, and South populations. The mitochondrial gene ND2 and the concatenated nuclear dataset also support these clades, whereas the individual nuclear gene trees were largely unresolved (Supplementary Figs. 2.1-2.6). In all well-resolved phylogenies, North and Central were recovered as sister clades. Generally speaking, our BEAST divergence time analysis estimated that the three clades diverged during the Pliocene and diversified during the Pleistocene (Fig. 2.2, Table 2.1).

Population genetic analyses

Phase reconstructed 13 unique haplotypes from 43 individuals for APP, 22 haplotypes for BRCA, 11 for Rag2, and 27 for vWF, although two individuals were not included in the vWF analysis due to short sequence length. Haplotype networks show the relationships among these haplotypes for each gene (Fig. 2.3).

Using the method of Evanno et al. (2005), K=3 was selected as the optimal number of populations within *Oryzorictes hova* in our STRUCTURE analysis (Supplementary Fig. 2.7). These population subdivisions correspond to the North, Central, and South clades identified in phylogenetic analyses (Fig. 2.2). All individuals were assigned to their respective populations with >90% confidence.

Values of F_{ST} among North, Central, and South clades were high, ranging from 0.74 – 0.85 (Table 2.2), and the AMOVA estimated that 80.24% of genetic variance was among clades (compared to 19.76% within clades; Table 2.3).

Extended Bayesian skyline plots for each of the North, Central, and South populations show a general trend of increasing population size from ~1 MYA – 250 KYA, followed by a population decline until present day (Fig. 2.4). In all plots, the 95% highest posterior density interval becomes wider (i.e., uncertainty in the value of theta increases) in the recent past (≤25 KYA).

Morphological analyses

ANOVA revealed that *O. hova* is significantly larger than *O. tetradactylus* in 26 of the 31 craniomandibular measurements, and that *O. tetradactylus* is significantly larger than *O. hova* in one measurement (LCP, defined in Supplementary Table 2.3). None of the measurements we evaluated were sexually dimorphic in either species (Supplementary Fig. 2.8).

For all multivariate results (PCA and CVA), there were no substantial differences between the available-case dataset and our imputed dataset; therefore, only the results of our imputed dataset, which included a greater number of variables and individuals, are shown in tables and figures, while the results of the available-case dataset analyses are provided in appendices (Supplementary Tables 2.4-2.7; Supplementary Figs. 2.9-2.11). Bivariate plots of the first two principal components showed clear morphological separation of *O. hova* from *O. tetradactylus*, but did not separate the North, Central, and South *O. hova* clades (Fig. 2.5 and Supplementary Fig. 2.9).

Species delimitation analyses

We identified 18 craniomandibular measurements that were significantly larger in the North *O. hova* clade than in either of the other clades, but did not identify any significant differences between the Central and South clades (Supplementary Fig. 2.10). In our CVA, specimens that were used to train the canonical variates (i.e., individuals sequenced as North, Central, or South) were clearly separated in morphospace (Fig. 2.6a and Supplementary Fig. 2.11) and were 100% correctly classified (Table 2.6 and Supplementary Table 2.6); however, many unsequenced specimens could not be confidently assigned to a clade (Supplementary Table 2.7), and were recovered as intermediate or extralimital to all three clades in morphospace. Believing this to be a result of the small size of our training dataset (i.e., the low number of individuals with sequence data compared to individuals with measurement data), we conducted a second CVA in which individuals were assigned to clades based on geography. Any individuals collected at or north of –15° latitude were used as training points for the North clade; individuals collected between –18° and –20.5° latitude were assigned to the Central clade; and individuals collected at or south of –20.5° latitude were assigned to the South clade. Individuals falling outside these bounds were not used in the training dataset, and all other analytical parameters remained the same. In this analysis, there was more

overlap between clades in morphospace (Fig. 2.6b) and the percentage of correct classification ranged from 86.1% in the North clade to 96.8% in the South clade (Table 2.6 and Supplementary Table 2.6).

Individuals that were consistently assigned to the same clade in all CV analyses (Supplementary Table 2.7) were used to predict the range extents of each clade (Fig. 2.7). The geographic range of the North clade extends from 14°S to 18.01°S latitude – much further south than our molecular data suggest. The Central clade ranges from 18.17°S to 21.29°S latitude, and the South clade ranges from 22.21°S to 42.57°S latitude.

Our BP&P analyses, which used molecular data only, recovered a species tree with the North and Central clades as sister taxa and recovered PP values of 1.0 for all three clades within *O. hova* across all prior distributions (Fig. 2.8a). As expected, randomized tip labeling resulted in very low PPs (≤0.02) for both divergence events across all prior settings.

When morphological data were considered independently and in combination with sequence data, iBPP again recovered PP values of 1.0 for all three putative species across all prior combinations (Fig. 2.8b,c). Unexpectedly, randomized tip labeling (simulating panmixia) did not always result in low speciation probabilities; PP values ranged from 0.0 to 1.0, depending on the randomization replicate and the prior distributions of Θ and τ. Overall, PP values from randomized tip analyses were lowest in the combined morphology + sequence data analysis, and when prior distributions were set to large population sizes and shallow divergence times.

2.5 Discussion

Phylogeography

We identified three distinct populations of *Oryzorictes hova* corresponding to the northern, central, and southern regions of the island (Fig. 2.2), a pattern that is roughly concordant with Madagascar's three highland regions (visible in Fig. 2.7): northern (Tsaratanana, Marojejy and associated massifs), central (Ankaratra, Andringitra, and associated massifs), and southern (the Anosyenne Mountains and associated southern extensions). Inhospitable habitat in the valleys between these regions may have limited gene flow during the Pliocene when the initial clade divergences occurred. Today, *O. hova* is known to occur at low elevations – even at sea level – and the South population's range extends approximately 175 km into the Central Highlands; thus, it seems unlikely that valleys are a current barrier to gene flow. Our Bayesian skyline plots also show a trend of increasing population sizes until approximately 200 – 250 KYA followed by a decline until the present (Fig. 2.4), supporting the hypothesis that populations have undergone past periods of expansion and contraction. It is not clear from our data how populations have responded to the most recent glacial cycles (e.g., since the Last Glacial Maximum) due to a large amount of uncertainty in our Bayesian skyline plots during this time frame.

Future work using larger genetic datasets – including loci that are informative at more recent time scales – will enable reconstruction of a more detailed demographic history. Future research using comparative phylogeographic methods may also reveal a common force driving divergence across multiple humid-forest-dwelling taxa.

Species delimitation

Both our molecular and morphological datasets support the existence of three species (i.e., independently evolving metapopulations *sensu* GLSC) currently ascribed to *O. hova*. Our STRUCTURE and BP&P analyses, which used only the molecular dataset, and our iBPP analyses, which included morphological data, all produced high support values for three distinct groups with little admixture. Additionally, our CVAs confidently assigned most individuals to clade based on morphology alone (Fig. 2.6 and Supplementary Fig. 2.11; Table 2.6 and Supplementary Tables 2.6 and 2.7). We also revealed two potential zones where clades might come into contact; namely, the North and Central clades may come into contact around 18°S latitude, and the Central and South clades may come into contact around 21.29 - 21.59°S, 47.43° - 47.50°E (Fig. 2.7). Regions of syntopy with no evidence of hybridization are strong additional support for reproductive isolation.

Assigning formal names to these three species is problematic for several reasons. First, the name *Oryzorictes talpoides* G. Grandidier and Petit 1930 [currently a junior synonym of *O. hova* (see Goodman et al. 1999)] has priority assuming that the holotype of *O. talpoides* (a vouchered specimen in the Muséum national d'Histoire naturelle, MNHN-ZM-MO 1985-1616) belongs to one of the three clades we identified. However, this holotype was collected near Marovoay in northwestern Madagascar, a locality that falls latitudinally midway between our molecular samplings of the North and the Central clades and ca. 200 km west of the published geographic range of *O. hova* (Fig. 2.1) in habitat predicted to be unsuitable for the species (Goodman et al. 2013). With respect to morphospace, the CVA places the holotype of *O. talpoides* either in the North clade or extralimital to all other measured specimens of *O. hova* (Fig. 2.6 and Supplementary Fig. 2.11), whereas PCA suggests intermediacy of this holotype between all measured *O. hova* and *O. tetradactylus* (Fig. 2.5 and Supplementary Fig. 2.9). While several other unsequenced but measured specimens of *O. hova* and *O. tetradactylus* likewise fall outside of the bounded ranges as defined by molecular results, our point here is that the holotype of *O. talpoides* cannot at present be confidently assigned to any of the clades that we here recognize, and there is evidence that it is craniometrically distinct from all three. Previous authors (Grandidier and Petit 1930; Genest and Petter 1975) invoked differences in relative pollex (thumb) length as a differentiator between *O. talpoides* and *O. hova*, but this has not been rigorously investigated and Goodman et al. (1999) attributed these perceived differences to intraspecific variation in *O. hova* sensu lato. Regardless, craniodental

morphology (this study), habitat differences (Goodman et al. 2013), and potential geographic isolation collectively suggest that *O. talpoides* may represent yet another species, in addition to and distinct from the three molecular clades of *O. hova* recovered in this study. However, we defer formal resurrection of *O. talpoides* pending resolution of additional outstanding uncertainties pertaining to type material of *Oryzorictes* (see below) that would be greatly aided by additional *Oryzorictes* material from this portion of the island.

Several problems surrounding the holotype of *O. hova* A. Grandidier, 1870 (MNHN-ZM-MO 1887-874) have also led to confusion about taxonomic priority as it pertains to assigning new species names. The type locality of *O. hova* was described as "Ankaye et Antsianak" (Grandidier 1870, p. 50). "Ankaye" likely refers to the Ankay Valley south of Lac Alaotra, while "Antsianak" to the Forêt Antsianaka (=Sihanaka) southeast of Lac Alaotra; the latter locality was referred to by Grandidier as "lac d'Antsianake" (Grandidier, 1872, Sibree, 1877). We estimate the geographic coordinates of this collection locality at approximately 18.42°S latitude and 48.75°E longitude (Fig. 2.1, Fig. 2.8), which falls near the projected contact zone between the North and Central clades. The *O. hova* holotype (MNHN-ZM-MO 1887-874) is consistently extralimital to all three clades in morphospace (Figs. 2.5 and 2.6; Supplementary Figs. 2.9 and 2.11), and is variably assigned either to the North or Central clade depending on the analysis (Supplementary Table 2.7). Two additional individuals were collected in 2011 from the same area (18.43°S, 48.79°E) and were included in our CVA: UADBA 33279 clustered with the Central clade, while UADBA 33306 clustered either with the North clade or the Central clade depending on the analysis (Supplementary Table 2.7). These results lend additional support to the idea that this locality represents a contact zone between the North and Central clades. Because we cannot be confident about the clade membership of the *O. hova* holotype, formal description of the other putative species must await DNA sequencing or other means of definitive group assignment to correctly apply the rules of taxonomic nomenclature.

Far more troubling and relevant to this discussion, however, is the persistent confusion regarding the purported type specimen (MNHN-ZM-MO 1887-874). Alfred Grandidier—an accomplished naturalist and astute observer—very clearly indicated the presence of three upper and three lower incisors in his description of the genus and species (Grandidier 1870), though he did not identify a holotype by number in the original description, a common omission in his era. However, as Goodman et al. (1999) and one of us (L.E.O.) have independently determined, not only are third upper and lower incisors entirely absent on the specimen labeled as the holotype, there is no indication of them having ever been present, and the premaxilla immediately behind the upper second incisor slopes sharply upward to create the deeply concave diastema that accommodates the crown of the lower canine during occlusion. The specimen in question is an adult female with fully erupted molars and permanent antemolars at all other dental loci,

although the cusps are relatively unworn. While it is not uncommon for one or more upper or lower teeth at a particular locus to be missing, broken, or worn completely down to the alveolar margin in some individuals, of the 269 *Oryzorictes* specimens coded at every dental locus by one of us (L.E.O.), only the specimen in question is completely missing the third incisor in all four quadrants. Likewise, of over 1,400 adult skulls of numerous species of *Microgale* (sister lineage to *Oryzorictes*) inspected by L.E.O., none are missing all four third incisors. The singularity of MNHN-ZM-MO 1887-874 (and its glaring conflict with the original description) is therefore striking. Equally puzzling are two fluid-preserved neonates assigned the same catalog number in the type collection of the Muséum national d'Histoire naturelle, Paris. If these are, indeed, the offspring of the holotype, it is hard to imagine that Grandidier (1870) would not have mentioned them, particularly in light of how little was known about tenrecid reproduction at the time. Based on these anomalies, we question the authenticity of the purported holotype of *O. hova* as the basis for Grandidier's (1870) description of the genus and species.

Performance of species validation methods

We validated species limits in *O. hova* using sequence data, craniomandibular measurement data, and both datasets together, and we tested the sensitivity of BP&P and iBPP to false positives by randomly assigning individuals to the North, Central, and South groups, which simulates a panmictic population and should result in low support for three species. The sequence-data-only BP&P analysis with randomized individual group assignments resulted in positive but very low PP values for three species, as expected. Disconcertingly, in the measurement-only iBPP analysis randomization of individual assignments did not always result in low PP values (values ranged from 0.0 to 1.0, Fig. 2.8). This is problematic because it suggests that 1.0 PP values, even in non-randomized analyses, could represent false positives. We cannot fully explain this result, but we note that the morphological dataset used in this analysis was very small as a result of the data reduction techniques that were used to meet the non-collinearity assumption of iBPP [although no smaller than the morphological datasets used with iBPP in other studies (e.g., Huang and Knowles 2016, Pyron et al. 2016)]. This result highlights the importance of using individual assignment randomization or other sensitivity analyses as a control for species validation methods. Importantly, when we used both measurement data and sequence data together, PP values for the simulated panmictic population decreased to 0.0 across all prior combinations. We, therefore, are in agreement with previous authors (e.g., Edwards and Knowles 2014) that the accuracy of species delimitation increases when multiple data types are used.

2.6 Conclusions

The molecular, morphological, and geographic evidence presented here supports the recognition of at least three distinct species currently referable to *Oryzorictes hova*. Assigning species names to these lineages will be the subject of future work, involving clarification of the confusion surrounding the holotype of *O. talpoides* and the putative holotype of *O. hova*. This study provides additional evidence for the latitudinal phylogeographic structure that has been identified in other humid-forest-dwelling taxa on Madagascar, and points to highland regions (possibly functioning as historical refugia) as a driving factor of population divergence. We echo other authors in advocating the use of multiple lines of evidence for species delimitation whenever possible (Carstens et al. 2013; Edwards and Knowles 2014), and we emphasize the need of continued collecting of voucher specimens on Madagascar, as these provide crucial information for documenting and interpreting biodiversity.

2.7 Acknowledgements

This work was supported by the National Science Foundation [grants numbers DEB-1120904, IOS-9623454, and a Graduate Research Fellowship], the Society of Systematic Biologists, the American Society of Mammalogists, the University of Alaska Museum, the WWF Madagascar, the Volkswagen Foundation, and an Alaska EPSCoR Undergraduate Research Fellowship. For access to tissues and voucher specimens under their care, we are deeply grateful to the late William T. Stanley, Bruce D. Patterson, Lawrence R. Heaney, Adam W. Ferguson, and Ben D. Marks (Field Museum of Natural History); Daniel Rakotondravony (Université d'Antananarivo, Mention Zoologie et Biodiversité Animale); Linda K. Gordon and Michael D. Carleton (National Museum of Natural History, Smithsonian Institution); Darrin P. Lunde, Christopher J. Raxworthy, Eileen Westwig, and Nancy B. Simmons (American Museum of Natural History); Roberto Portela Miguez, Paula D. Jenkins, and Louise Tomsett (Natural History Museum, London); Mark Omura and Judy Chupasko (Museum of Comparative Zoology, Harvard); Cécile Callou, Christiane Denys, and Violaine Colin (Muséum national d'Histoire naturelle, Paris); and James L. Patton and Christopher J. Conroy (Museum of Vertebrate Zoology, University of California, Berkeley). We thank Michi Schulenberg for her outstanding preparation of specimens at the Field Museum. Finally, we thank J. Peters, A. Moore, S. M. Powers, J. L. Fiely, H. C. Lanier, E. Humphries, and T. E. Roberts for their invaluable advice and assistance with lab work and analyses. Sushma Reddy and an anonymous reviewer provided thoughtful and thorough feedback that greatly improved the final manuscript.

2.8 References

Burns, S.J., Godfrey, L.R., Faina, P., McGee, D., Hardt, B., Ranivoharimanana, L., Randrianasy, J., 2016. Rapid human-induced landscape transformation in Madagascar at the end of the first millennium of the Common Era. Quat. Sci. Rev. 134, 92-99.

Carstens, B.C., Pelletier, T.A., Reid, N.M., Satler, J.D., 2013. How to fail at species delimitation. Mol. Ecol. 22, 4369-4383.

Clement, M., Posada, D., Crandall, K., 2000. TCS: A computer program to estimate gene genealogies. Mol. Ecol. 9, 1657-1660.

Coyne, J.A., Orr, H.A., 2004. Speciation. Sinauer Associations, Sunderland, MA, pp. 83-124.

de Queiroz, K., 2007. Species concepts and species delimitation. Syst. Biol. 56, 879-886. Drummond,

A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Dufils, J.M., 2003. Remaining forest cover. In: Goodman, S.M., Benstead, J.P. (Ed.), The Natural History of Madagascar. The University of Chicago Press, Chicago, pp. 88-96.

Earl, D.A., vonHoldt, B.M., 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4, 359-361.

Edwards, D.L., Knowles, L.L., 2014. Species detection and individual assignment in species delimitation: Can integrative data increase efficacy? Proc. R. Soc. Lond. B: Biol. Sci. 281, 20132765.

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Mol. Ecol. 14, 2611-2620.

Everson, K.M., Soarimalala, V., Goodman, S.M., Olson, L.E., 2016. Multiple loci and complete taxonomic sampling resolve the phylogeny and biogeographic history of tenrecs (Mammalia: Tenrecidae) and reveal higher speciation rates in Madagascar's humid forests. Syst. Biol. 65, 890-909.

Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Res. 10, 564-567.

Flynn, J.J., Wyss, A.R., 2003. Mesozoic terrestrial vertebrate faunas: The early history of Madagascar's vertebrate diversity. In: Goodman, S.M., Benstead, J.P. (Ed.), The Natural History of Madagascar. The University of Chicago Press, Chicago, pp. 34-40.

Gamisch, A., Fischer, G.A., Comes, H.P., 2016. Frequent but asymmetric niche shifts in *Bulbophyllum* orchids support environmental and climatic instability in Madagascar over Quaternary time scales. BMC Evol. Biol. 16, 1.

Gehring, P.S., Pabijan, M., Randrianirina, J.E., Glaw, F., Vences, M., 2012. The influence of riverine barriers on phylogeographic patterns of Malagasy reed frogs (*Heterixalus*). Mol. Phylogenet. Evol. 64, 618-632.

Genest, H., Petter, F., 1975. Part 1.1. Family Tenrecidae. In: Meester, J., Setzer, H.W. (Ed.), The mammals of Africa: An identification manual. Smithsonian Institution Press, Washington, pp. 1-7.

Goodman, S.M., 2003a. Introduction to the mammals. In: Goodman, S.M., Benstead, J.P. (Ed.), The Natural History of Madagascar. The University of Chicago Press, Chicago, pp. 1159-1186.

Goodman, S.M., 2003b. *Oryzorictes*, mole tenrec or rice tenrec. In: Goodman, S.M., Benstead, J.P. (Ed.), The Natural History of Madagascar. The University of Chicago Press, Chicago, pp. 1278-1281.

Goodman, S.M., Benstead, J.P., 2005. Updated estimates of biotic diversity and endemism for Madagascar. Oryx 39, 73-77.

Goodman, S.M., Jenkins, P.D., Pidgeon, M., 1999. The Lipotyphla (Tenrecidae and Soricidae) of the Reserve Naturelle Integrale d'Andohahela, Madagascar. Fieldiana Zool. n.s. 94, 187-216.

Goodman, S.M., Rakotondravony, D. 2000. The effects of forest fragmentation and isolation on insectivorous small mammals (Lipotyphla) on the Central High Plateau of Madagascar. J. Zool. 250, 193-200.

Goodman, S.M., Rakotondravony, D., Randriamanantsoa, H.N., Rakotomalala-Razanahoera, M., 2005. A new species of rodent from the montane forest of central eastern Madagascar (Muridae: Nesomyinae: *Voalavo*). Proc. Biol. Soc. Wash. 118, 863-873.

Goodman, S.M., Raxworthy, C.J., Maminirina, C.P., Olson., L.E., 2006. A new species of shrew tenrec (*Microgale jobihely*) from northern Madagascar. J. Zool. 270, 384-398.

Goodman, S.M., Soarimalala, V., Raheriarisena, M., Rakotondravony, D., 2013. Small mammals or tenrecs (Tenrecidae) and rodents (Nesomyidae). In: Goodman, S.M., Raherilalao, M. (Ed.), Atlas of Selected Land Vertebrates of Madagascar. Association Vahatra, Antananarivo, Madagascar, pp. 211-269.

Grandidier, A., 1870. Description de quelques animaux nouveaux découverts à Madagascar en Novembre 1869. Rev. Mag. Zool. Pur. App. 22, 50.

Grandidier, A., 1872. Rapports sur une mission ä Madagascar. Archives des Missions Scientifiques 7, 445-477.

Grandidier, G., Petit, G., 1930. Description d'une espèce nouvelle d'insectivore malgache, suivie de remarques critiques sur le genre *Oryzoryctes* [sic]. Bull. Mus. Hist. Nat. 2, 498.

Guillot, G., Renaud, S., Ledevin, R., Michaux, J., Claude, J., 2012. A unifying model for the analysis of phenotypic, genetic, and geographic data. Syst. Biol. 61, 897-911.

Harper, G.J., Steininger, M.K., Tucker, C.J., Juhn, D., Hawkins, F., 2007. Fifty years of deforestation and forest fragmentation in Madagascar. Environ. Conserv. 34, 325-333.

Heled, J., Drummond, A.J., 2010. Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. 27, 570-580.

Huang, J., Knowles, L.L., 2016. The species versus subspecies conundrum: Quantitative delimitation from integrating multiple data types within a single Bayesian approach in hercules beetles. Syst. Biol. 65, 685-699.

Huelsenbeck, J.P., Andolfatto, P., Huelsenbeck, E.T., 2011. Structurama: Bayesian inference of population structure. Evol. Bioinform. 7, 55.

IUCN Species Survival Commission, 2012. IUCN Red List categories and criteria: Version 3.1 second edition. IUCN, Gland, Switzerland, pp. 10-22.

Jakobsson, M., Rosenberg, N.A., 2007. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23, 1801-1806.

Lanfear, R., Calcott, B., Ho, S.Y., Guindon, S., 2012. PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29, 1695-1701.

Leaché, A.D., Fujita, M.K., 2010. Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). Proc. R. Soc. Lond. B: Biol. Sci. 662, 1-7.

Lehtinen, R.M., Nussbaum, R.A., Richards, C.M., Cannatella, D.C., Vences, M., 2007. Mitochondrial genes reveal cryptic diversity in plant-breeding frogs from Madagascar (Anura, Mantellidae, *Guibemantis*). Mol. Phylogenet. Evol. 44, 1121-1129.

Milne-Edwards, A., Grandidier, A., 1882. Description d'une nouvelle espèce d'insectivore de Madagascar. Le Naturaliste 4, 349.

Mittermeier, R.A., Louis Jr., E.E., Langrand, O., Schwitzer, C., Gauthier, C.A., Rylands, A.B., Rajaobelina, S., Ratsimbazafy, J., Rasoloarison, R., Hawkins, F., Roos, C., Richardson, M., Kappeler, P.M., 2015. Lémuriens de Madagascar. Publications scientifiques du Muséum national d'Histoire naturelle, Paris.

O'Meara, B.C., 2010. New heuristics methods for joint species tree inference and species delimitation. Syst. Biol. 59, 59-73.

Olivieri, G., Zimmermann, E., Randrianambinina, B., Rasoloharijaona, S., Rakotondravony, D., Guschanski, K., Radespiel, U., 2007. The ever-increasing diversity in mouse lemurs: Three new species in north and northwestern Madagascar. Mol. Phylogenet. Evol. 43, 309-327.

Olson, L.E., 2013. Tenrecs. Curr. Biol. 23, R5-R8.

Olson, L.E., Goodman, S.M., 2003. Phylogeny and biogeography of tenrecs. In: Goodman, S.M., Benstead, J.P. (Ed.), The Natural History of Madagascar. The University of Chicago Press, Chicago, pp. 1235-1242.

Olson, L.E., Goodman, S.M., Yoder, A.D., 2004. Illumination of cryptic species boundaries in long-tailed shrew tenrecs (Mammalia: Tenrecidae; *Microgale*): New insights into geographic variation and distributional constraints. Biol. J. Linn. Soc. 83, 1-22.

Olson, L.E., Rakotomalala, Z., Hildebrandt, K.B.P., Lanier, H.C., Raxworthy, C.J., Goodman, S.M., 2009. Phylogeography of *Microgale brevicaudata* (Tenrecidae) and description of a new species from western Madagascar. *J. Mamm.* 90, 1095-1110.

Pastorini, J., Thalmann, U., Martin, R.D., 2003. A molecular approach to comparative phylogeography of extant Malagasy lemurs. Proc. Natl. Acad. Sci. 100, 5879-5884.

Perl, R.G.B., Nagy, Z.T., Sonet, G., Glaw, F., Wollenberg, K.C., Vences, M., 2014. DNA barcoding Madagascar's amphibian fauna. Amphibia-Reptilia 35, 197-206.

Posada, D., Crandall, K.A., 1998. Modeltest: Testing the model of DNA substitution. Bioinformatics 14, 817-818.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945-959.

Poux C., Madsen, O., Marquard, E., Vieites, D.R., DeJong, W.W., Vences, M., 2005. Asynchronous colonization of Madagascar by the four endemic clades of primates, tenrecs, carnivores, and rodents as inferred from nuclear genes. Syst. Biol. 54, 719-730.

Pyron, R.A., Hsieh, F.W., Lemmon, A.R., Lemmon, E.M., Hendry, C.R., 2016. Integrating phylogenomic and morphological data to assess candidate species-delimitation models in brown and red-bellied snakes (*Storeria*). Zool. J. Linn. Soc. 177, 937-949.

Quéméré, E., Amelot, X., Pierson, J., Crouau-Roy, B., Chikhi, L., 2012. Genetic data suggest a natural prehuman origin of open habitats in northern Madagascar and question the deforestation narrative in this region. Proc. Natl. Acad. Sci. 109, 13028-13033.

Rosenberg, N.A., 2003. DISTRUCT: A program for the graphical display of population structure. Mol. Ecol. Notes 4, 137-138.

Sibree, J., 1877. The Sihanaka and their country. Antananarivo Annual and Madagascar Magazine 3, 51-69.

Silva, A.P.D., 2011. Two-group classification with high-dimensional correlated data: A factor model approach. Comput. Stat. Data An. 55, 2975-2990.

Solís-Lemus, C., Knowles, L.L., Ané, C., 2014. Bayesian species delimitation combining multiple genes and traits in a unified framework. Evolution 69, 492-507.

Stanhope, M.J., Waddell, V.G., Madsen, O., De Jong, W., Hedges, S.B., Cleven, G.C., Kao, D., Springer, M.S., 1998. Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. Proc. Natl. Acad. Sci. USA 95, 9967-9972.

Stephens, M., Donnelly, P., 2003. A comparison of Bayesian methods for haplotype reconstruction. Am. J. Hum. Genet. 73, 1162-1169.

Stephens, M., Smith, N.J., Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. 68, 978-989.

Stephenson, P.J., Soarimalala, V., Goodman, S., 2016. *Oryzorictes hova*. The IUCN Red List of Threatened Species 2016: e.T40589A97203050. doi: 10.2305/IUCN.UK.2016-1.RLTS.T40589A97203050.en. Downloaded on 14 September 2017.

Stephenson, P.J., Soarimalala, V., Goodman, S., 2016. *Oryzorictes tetradactylus*. The IUCN Red List of Threatened Species 2016: e.T40589A97203050. doi: 10.2305/IUCN.UK.2016-1.RLTS.T40589A97203050.en. Downloaded on 14 September 2017.

Swofford, D.L., 2002. PAUP*: Phylogenetic Analyses Using Parsimony (*and other methods). Sinauer Associates, Sunderland, MA.

van Buuren, S., Groothuis-Oudshoorn, K., 2011. mice: Multivariate Imputation by Chained Equations in R. J. Stat. Softw. 45, 1-67.

Veron, G., Dupré, D., Jennings, A.P., Gardner, C.J., Hassanin, A., Goodman, S.M., 2017. New insights into the systematics of Malagasy mongoose-like carnivorans (Carnivora, Eupleridae, Galidiinae) based on mitochondrial and nuclear DNA sequences. J. Zool. Syst. Evol. Res. 55, 250-264.

Veron, G., Goodman, S.M., 2017. One or two species of the rare Malagasy carnivoran *Eupleres* (Eupleridae)? New insights from molecular data. Mammalia (preprint). doi: 10.1515/mammalia-2016-0182. Downloaded on 28 September 2017.

Vieites, D.R., Chiarai, Y., Vences, M., Androeone, F., Rabemanajara, F., Bora, P., Nieto-Roman, S., Meyer, A., 2006. Mitochondrial evidence for distinct phylogeographic units in the endangered Malagasy poison frog *Mantella bernhardi*. Mol. Ecol. 15, 1617-1625.

Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution 38, 1358-1370.

Wickham, H., 2009. ggplot2: Elegant graphics for data analysis. Springer-Verlag, New York.

Wilmé, L., Goodman, S.M., Ganzhorn, J.U., 2006. Biogeographic evolution of Madagascar's microendemic biota. Science 312, 1063-1065.

Wirta, H., 2009. Complex phylogeographical patterns, introgression and cryptic species in a lineage of Malagasy dung beetles (Coleoptera: Scarabaeidae). Biol. J. Linn. Soc. 96, 942-955.

Yang, Z., Rannala, B., 2010. Bayesian species delimitation using multilocus sequence data. Proc. Natl. Acad. Sci. 107, 9264-9269.

Yang, Z., Rannala, B., 2014. Unguided species delimitation using DNA sequence data from multiple loci. Mol. Biol. Evol. 31, 3125-3135.

Yoder, A.D., 2007. Lemurs: A quick guide. Curr. Biol. 17, 866-868.

Yoder, A.D., Burns, M.M., Zehr, S., Delefosse, T., Veron, G., Goodman, S.M., Flynn, J.J., 2003. Single origin of Malagasy Carnivora from an African ancestor. Nature 421, 734-737.

Yoder, A.D., Cartmill, M., Ruvolo, M., Smith, K., Vilgalys, R., 1996. Ancient single origin for Malagasy primates. Proc. Natl. Acad. Sci. USA 93, 5122-5126.

Yoder, A.D., Nowak, M.D., 2006. Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. Ann. Rev. Ecol. Syst. 37, 405-431.

Zwickl, D.J., 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion (PhD). Ph.D. Dissertation, University of Texas at Austin, Austin, Texas, USA.

2.9 Tables

Table 2.1. Estimated divergence times for all clade splits within *Oryzorictes* and crown diversifications. HPD: highest posterior density interval.

Divergence event	Age (MYA)	95% HPD
O. hova – *O. tetradactylus* divergence	5.13	3.67 – 6.63
(North + Central) – South clade divergence	4.03	2.75 – 5.50
North – Central clade divergence	2.82	1.72 – 4.25
Crown diversification of North clade	0.93	0.49 – 1.57
Crown diversification of Central clade	0.97	0.49 – 1.62
Crown diversification of South clade	1.21	0.65 – 2.05

Table 2.2. Pairwise FST values among *O. hova* populations. Permutation test found that all values are significant ($p < 0.001$).

	North	Central	South
North	-	0.73612	0.85192
Central		-	0.78063
South			-

Table 2.3. AMOVA results within the three recognized *O. hova* clades. Populations were defined by the groups recovered in the STRUCTURE analysis. d.f. = degrees of freedom.

	d.f.	Sum of Squares	Variation	Significance
Among Populations	2	1480.711	80.24%	P<0.00001
Within Populations	83	538.277	19.76%	P<0.00001
Total	85	2018.988		

Table 2.4. Variable loadings on the first two principal components using the imputed morphometric dataset. Variable loadings from the PCA conducted with the available-case dataset are provided in Table 2.4.

	PC1	PC2
Ai1	-0.2030523	0.14279286
AC	-0.1456294	-0.1066016
BB	-0.180674	-0.16701517
C1W	-0.1688922	-0.20288301
CIL	-0.2015641	0.1739158
CCor	-0.1806463	-0.0728133
ML	-0.2112162	0.09990373
CC1	-0.205457	0.14430087
CI2	-0.2039978	0.160729
CP2	-0.2016806	0.14670181
CP3	-0.2002432	0.14446757
CPM	-0.1932232	0.20310756
MH	-0.1936926	-0.06125388
I2W	-0.172856	-0.17461484
LCP	-0.1178013	0.32597398
M1WP	-0.1812882	-0.22347734
M2WA	-0.1834071	-0.2072399
M2WP	-0.1759932	-0.23196322
M3W	-0.183246	-0.18842961
m3c1	-0.1982706	0.05788386
m3i1	-0.1881504	0.13121026
MC2	-0.1178822	-0.15916793
NC	-0.1835008	0.22634008
ZN	-0.1376954	0.27203627
NAW	-0.1661102	-0.22409943
P2W	-0.173915	-0.10300999
P3W	-0.1347112	-0.13642592
P4WP	-0.1837247	-0.21110253
PZ	-0.1651997	0.21368864
UTR	-0.1952807	0.11316546
ZB	-0.1668801	-0.2168952
Cumulative %	65.24%	82.89%

Table 2.5. Variable loadings on the first two canonical variates for two analyses: the first analysis in which the training dataset included only individuals with DNA sequence data, and the second analysis in which the training dataset included individuals that were assigned by geographic locality. Both analyses used the imputed morphometric dataset; variable loadings from analyses using the available-case morphometric dataset are provided in Supplementary Table 2.6. Variables that are not shown were removed during the shrunken covariance procedure. In both analyses, the first two canonical variates explain 100% of the variance.

CVA loadings: training data assigned by sequencing (Fig. 2.6a)		
	CV1	**CV2**
Ai1	6.738	5.636
CIL	-29.212	-8.102
CCor	-0.912	2.240
ML	-0.489	-2.191
CC1	3.100	7.060
CI2	31.095	19.986
CP2	-9.648	-0.719
CP3	1.130	-14.140
CPM	-5.157	-20.950
LCP	-1.731	3.988
M1WP	-0.904	7.041
M2WA	-13.948	-12.026
M2WP	10.362	6.972
m3c1	-25.490	0.695
m3i1	14.799	-4.296
NC	4.952	6.768
ZN	-9.825	-7.416
PZ	7.821	6.668
UTR	5.182	4.427

CVA loadings: training data assigned by geography (Fig. 2.6b)		
	CV1	**CV2**
Ai1	-3.506	-1.813
AC	0.795	2.121
BB	0.919	-1.043
C1W	1.905	1.981
CIL	-4.546	-5.605
CCor	0.702	-0.716
ML	0.458	0.678
CC1	-2.046	3.742
CI2	1.707	2.639
CP2	-0.752	-3.805
CP3	2.831	1.236
CPM	6.443	1.771
MH	-1.458	1.576
I2W	-1.958	-6.563
LCP	-1.990	1.047
M1WP	-2.627	0.513
M2WA	1.829	-4.107
M2WP	-1.873	0.672
M3W	-0.620	0.714
m3c1	-8.092	0.820
m3i1	5.783	-1.159
NC	-0.710	-1.880
ZN	0.575	1.612
NAW	-2.105	0.968
P2W	-2.262	-2.190
P4WP	1.845	2.293
PZ	0.301	-4.912
UTR	2.593	6.120
ZB	-0.038	1.493

Table 2.6. Classification matrices for two canonical variate analyses: the first analysis in which the training dataset included only individuals with DNA sequence data, and the second analysis in which the training dataset included individuals that were assigned by geographic locality. Both analyses used the imputed morphometric dataset; classification matrices from analyses using the available-case morphometric dataset are provided in Supplementary Table 2.7.

Classification matrix for CVA where training data are assigned by sequencing (Fig. 2.6a)				
	North	**Central**	**South**	**% Correct**
North	9	0	0	100%
Central	0	7	0	100%
South	0	0	13	100%
Classification matrix for CVA where training data are assigned by geography (Fig. 2.6b)				
	North	**Central**	**South**	**% Correct**
North	21	3	0	87.5%
Central	1	31	4	86.1%
South	0	1	30	96.8%

2.10 Figures

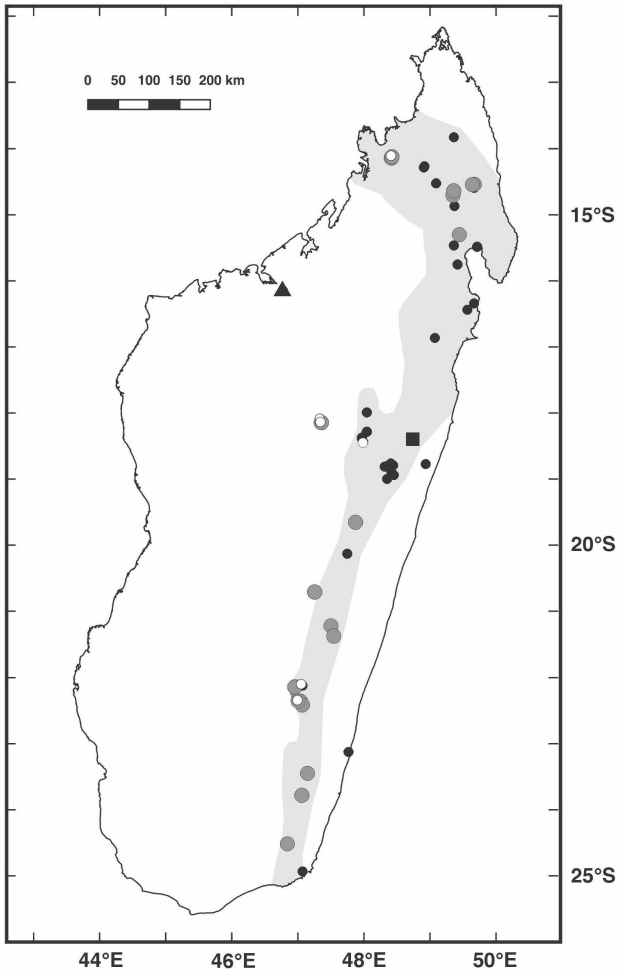


Figure 2.1. Map of Madagascar showing the range of *Oryzorictes hova* (shaded region) and the collection localities for specimens used in this study. Specimens are associated with sequence data and morphology data (gray circles), sequence data alone (white), or morphology data alone (black). Type localities of *O. hova* A. Grandidier, 1870 (square) and *O. talpoides* G. Grandidier and Petit, 1930 (triangle) are also shown. Range map was produced by the International Union for Conservation of Nature (Stephenson et al. 2016a) using data from Goodman et al (2013).

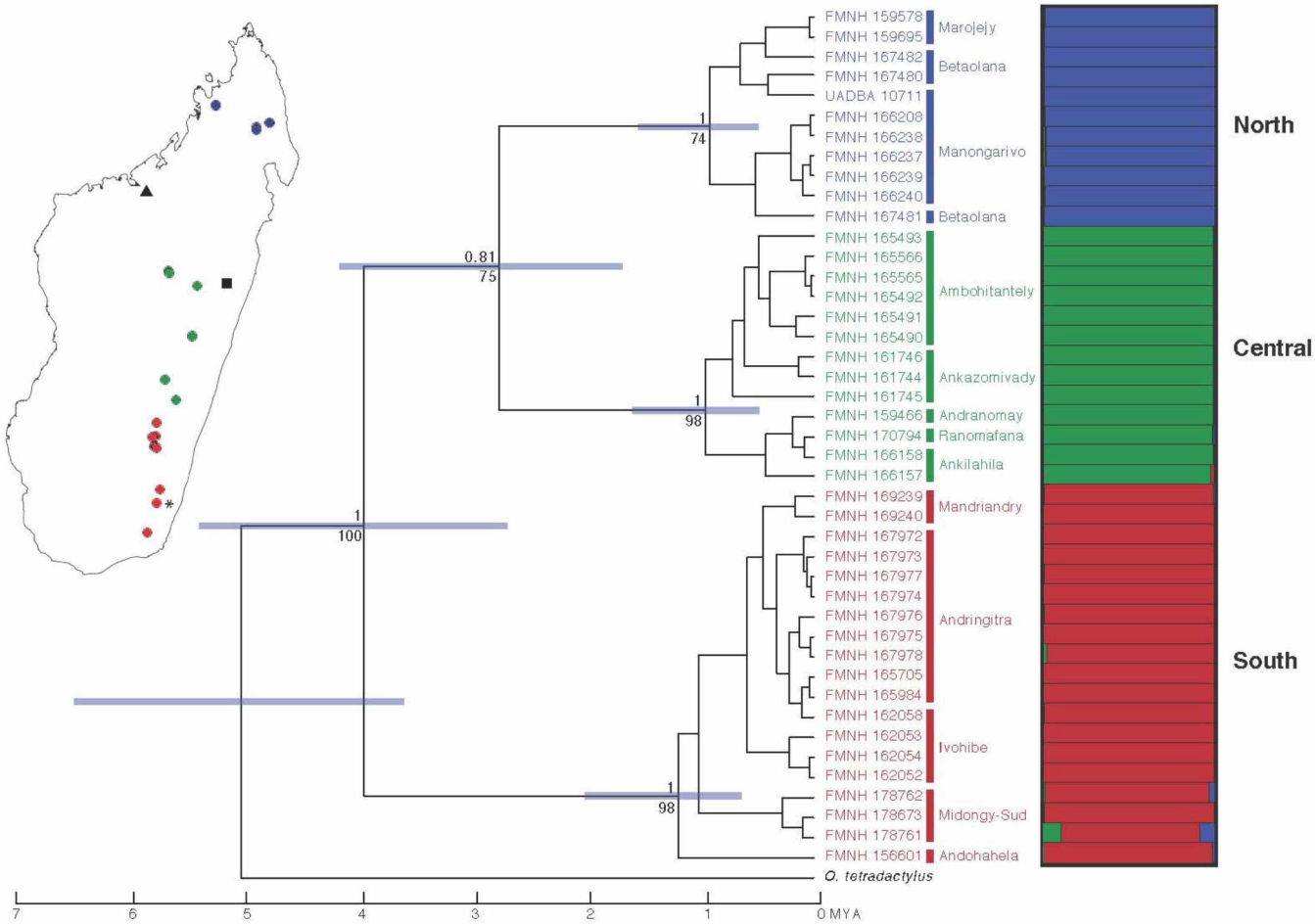


Figure 2.2. Phylogeny showing relationships and divergence times among *O. hova* individuals as determined from the five-gene concatenated BEAST analysis. Bayesian posterior probabilities (BPP) are shown above nodes and bootstrap support (BS) values from the concatenated maximum-likelihood analysis (Garli) are shown below. For clarity, only nodal support values for clade divergences and crown diversification events are shown. Purple nodal bars indicate the 95% highest posterior density intervals for the ages of important divergence events (listed in Table 2.1). Results from the STRUCTURE analysis are shown to the right of the phylogeny. Each horizontal bar represents an individual and each color represents the relative probability of membership in one of three main groups. Colors correspond to the North (blue), Central (green), and South (red) clades, which are also shown on the map (left). Collection localities are shown to the right of the bars after each tip label, with full locality descriptions provided in Supplementary Table 2.1. Type localities of *O. hova* (square) and *O. talpoides* (triangle) are also shown, although these were not included in molecular analyses. The asterisk on the map shows the collection locality of the single individual (FMNH 178761) with >5% admixture.

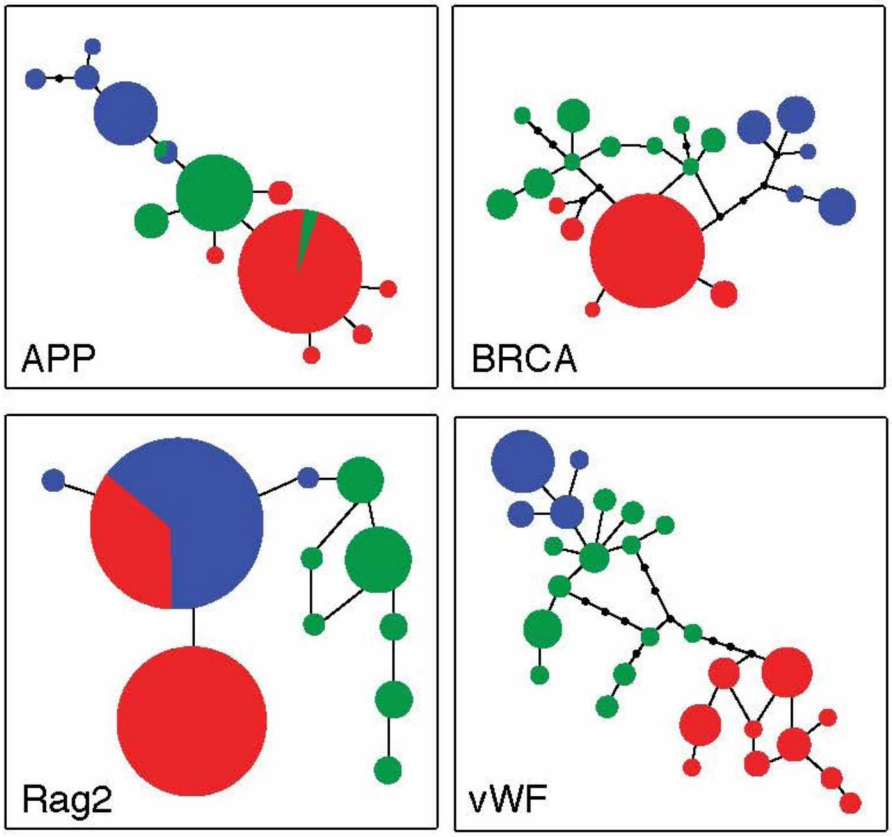


Figure 2.3. Networks from the recovered nuclear haplotypes in APP, BRCA, Rag2, and vWF for *O. hova*. Colors correspond to the North (blue), Central (green), and South (red) clades (see Fig. 2.2). The sizes of each circle correspond to the number of individuals with that haplotype. Small black dots represent unsurveyed haplotypes.

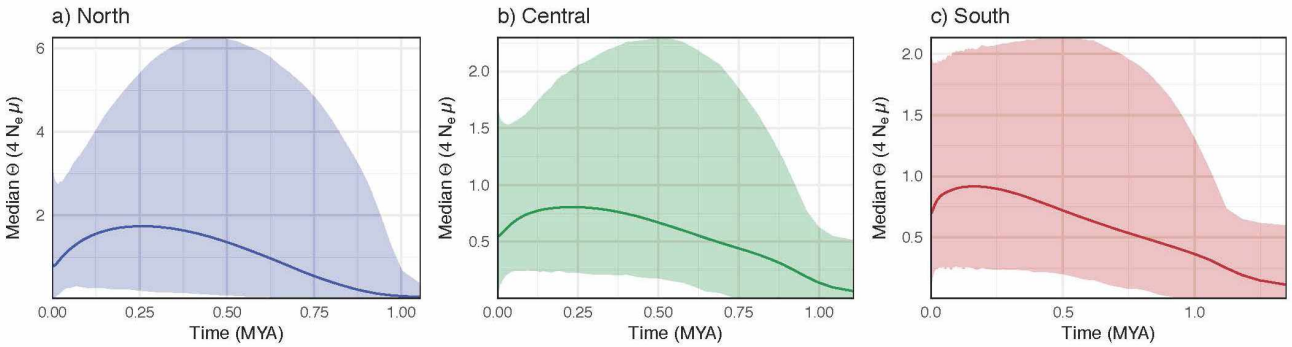


Figure 2.4. Extended Bayesian skyline plots for each of the three *O. hova* clades identified in this study. The solid line indicates the median estimate for theta (= 4Neμ) through time, while the lighter colored band shows the 95% highest posterior density interval.

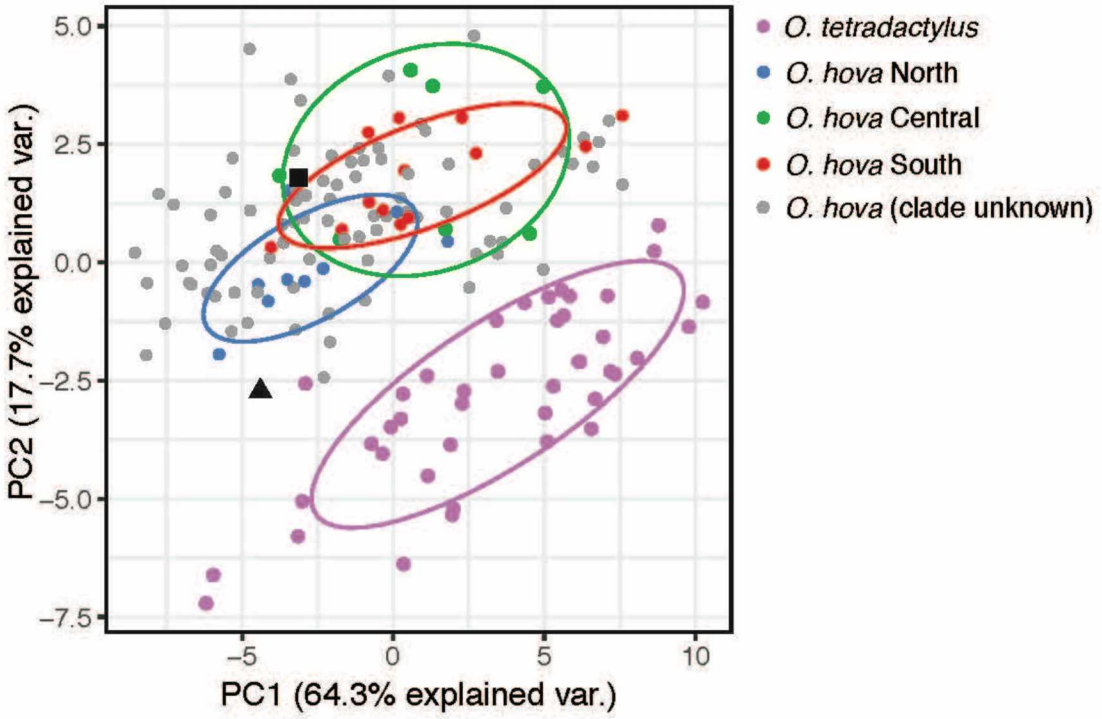


Figure 2.5. Bivariate plots of the first two principal components using the complete, imputed craniomandibular dataset. Results from analyses using the available-case dataset are provided in Fig. A.10. Individual points and 95% confidence ellipses are colored according to assignment to the North (blue), Central (green), or South (red) clade. Type specimens of *O. hova* and *O. talpoides* are indicated by a square and a triangle, respectively.

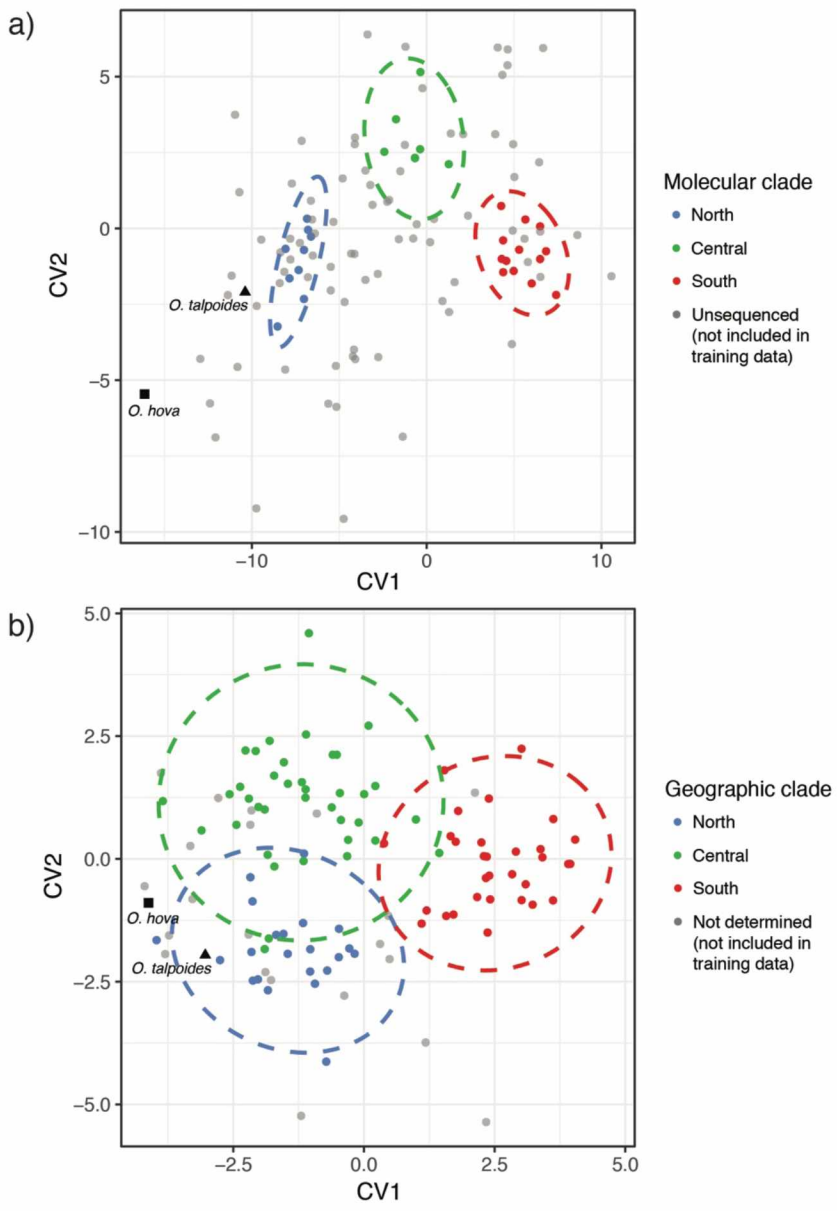


Figure 2.6. Bivariate plots of the first two canonical variates for the analysis in which (a) sequenced individuals were used to train the canonical variates, and (b) canonical variates were trained based on the geographic locations of individuals. The imputed dataset was used in both analyses; results from analyses using the available-case dataset are provided in Fig. A.11. Individual points and 95% confidence ellipses are colored according to assignment to the North (blue), Central (green), or South (red) clade. Type specimens of *O. hova* and *O. talpoides* are indicated by a square and a triangle, respectively.

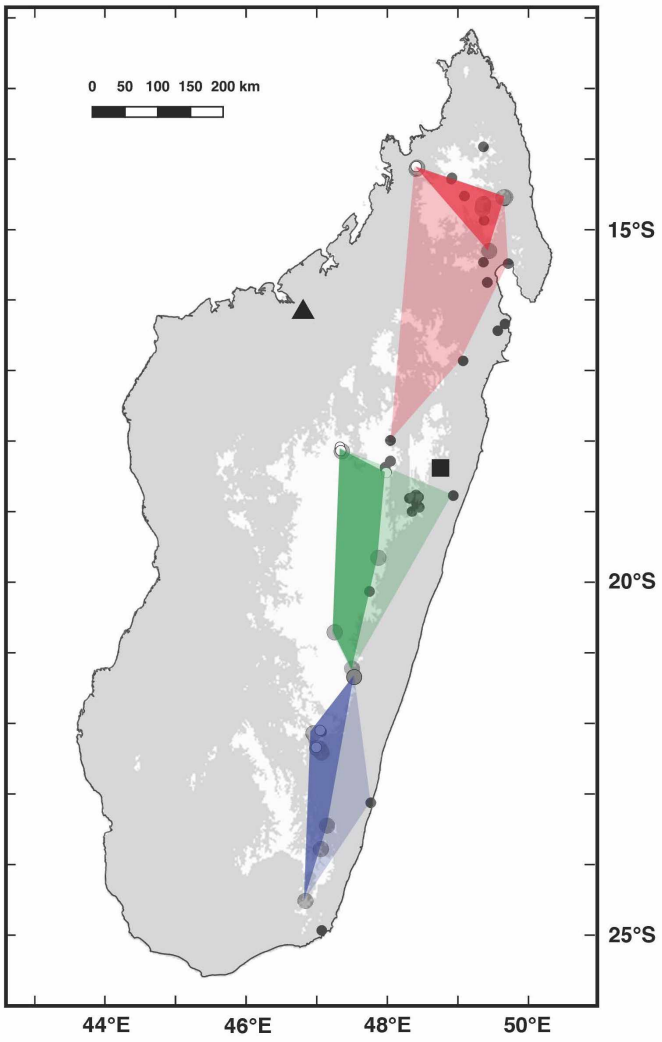


Figure 2.7. Map of Madagascar showing expected range extents (drawn as minimum convex polygons) for each of the three new candidate species of *Oryzorictes*: North (blue), Central (green), and South (red). Darker colors show the extents using only genetically confirmed specimens, while lighter colors expand on these regions by showing inferred clade memberships from our CVAs. White areas on the map show high-elevation regions (>900 m). Points are colored as in Fig. 1. Not all points could be confidently assigned to one of the three clades by CVA. Type localities of *O. hova* (square) and *O. talpoides* (triangle) are also shown.

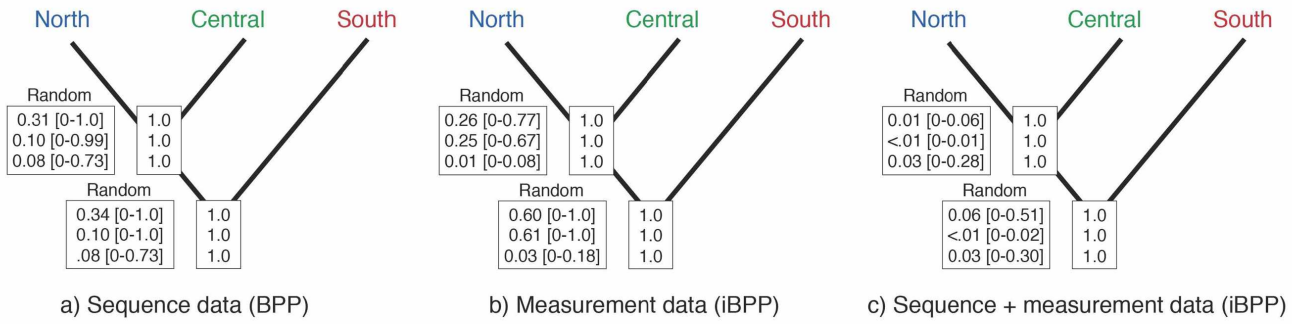


Figure 2.8. Bayesian species delimitation results from analyses that included (a) only sequence data, (b) only morphological data, or (c) both datasets together. The speciation probabilities are provided at each node for each combination of priors for θ and τ: top, θ = G(2,2000), τ = G(1,10); middle, θ = G(1,10), τ = G(1,10); bottom, θ = G(2,2000), τ = G(2,2000). Speciation probabilities when terminal labels are randomized are also provided to the left of each node, with the probability range provided in brackets.

2.11 Supplementary Tables

Supplementary Table 2.1. Museum catalog numbers, localities, and GenBank accession numbers for specimens associated with sequence data. Collection abbreviations are as follows: Field Museum of Natural History (FMNH), Université d'Antananarivo, Mention Zoologie et Biodiversité Animale (formerly Département de Biologie Animale, UADBA).

Collection	Catalog #	Latitude	Longitude	APP	BRCA1	ND2	RAG2	vWF	Locality	Locality Description
O. tetradactylus										
FMNH	156226	-14.0217	48.4183	KX015580	KX015448	KX015624	KX015492	KX015536	Andringitra	Réserve Naturelle Intégrale d'Andringitra, Cuvette de Pic Boby
O. hova, north clade										
FMNH	159578	-14.4400	49.7417	KX015581	KX015449	KX015625	KX015493	KX015537	Marojejy	Réserve Naturelle Intégrale de Marojejy, 10.5 km NW Manantenina
FMNH	159695	-14.4367	49.7750	KX015582	KX015450	KX015626	KX015494	KX015538	Marojejy	Réserve Naturelle Intégrale de Marojejy, 8 km NW Manantenina, along tributary of Manantenina River
FMNH	166208	-14.0217	48.4183	KX015586	KX015454	KX015630	KX015498	KX015542	Manongarivo	Réserve Spéciale de Manongarivo, 17.3 km SW Antanambao
FMNH	166237	-14.0000	48.4283	KX015588	KX015456	KX015632	KX015500	KX015544	Manongarivo	Réserve Spéciale de Manongarivo, 14.5 km SW Antanambao
FMNH	166238	-14.0000	48.4283	KX015587	KX015455	KX015631	KX015499	KX015543	Manongarivo	Réserve Spéciale de Manongarivo, 14.5 km SW Antanambao
FMNH	166239	-14.0000	48.4283	KX015589	KX015457	KX015633	KX015501	KX015545	Manongarivo	Réserve Spéciale de Manongarivo, 14.5 km SW Antanambao
FMNH	166240	-14.0000	48.4283	KX015590	KX015458	KX015634	KX015502	KX015546	Manongarivo	Réserve Spéciale de Manongarivo, 14.5 km SW Antanambao
FMNH	167480	-14.5383	49.4383	KX015584	KX015452	KX015628	KX015496	KX015540	Betaolana	Forêt de Betaolana, along Ambolokopatrika River, 8.5 km NW Ambodiangezoka
FMNH	167481	-14.5383	49.4383	KX015591	KX015459	KX015635	KX015503	KX015547	Betaolana	Forêt de Betaolana, along Ambolokopatrika River, 8.5 km NW Ambodiangezoka
FMNH	167482	-14.6100	49.4250	KX015583	KX015451	KX015627	KX015495	KX015539	Betaolana	Forêt de Betaolana, 11 km NW Ambodiangezoka
UADBA	10711	-14.0217	48.4183	KX015585	KX015453	KX015629	KX015497	KX015541	Manongarivo	Réserve Spéciale de Manongarivo, 17.3 km SW Antanambao
O. hova, central clade										
FMNH	159466	-18.4800	47.9550	KX015592	KX015460	KX015636	KX015504	KX015548	Andranomay	2 km NNE Andranomay, 13 km SE Anjozorobe
FMNH	161744	-20.7750	47.1683	KX015596	KX015464	KX015640	KX015508	KX015552	Ankazomivady	Forêt d'Ankazomivady, 28 km SSW Ambositra, 5 km SW Ambalamanakana
FMNH	161745	-20.7750	47.1683	KX015604	KX015472	KX015648	KX015516	KX015560	Ankazomivady	Forêt d'Ankazomivady, 28 km SSW Ambositra, 5 km SW Ambalamanakana
FMNH	161746	-20.7750	47.1683	KX015597	KX015465	KX015641	KX015509	KX015553	Ankazomivady	Forêt d'Ankazomivady, 28 km SSW Ambositra, 5 km SW Ambalamanakana

Supplementary Table 2.1 (continued).

Collection	Catalog #	Latitude	Longitude	APP	BRCA1	ND2	RAG2	vWF	Locality	Locality Description
FMNH	165490	-18.1683	47.2767	KX015598	KX015466	KX015642	KX015510	KX015554	Ambohitantely	Réserve Spéciale d'Ambohitantely, 24 km NE Ankazobe
FMNH	165491	-18.1683	47.2767	KX015599	KX015467	KX015643	KX015511	KX015555	Ambohitantely	Réserve Spéciale d'Ambohitantely, 24 km NE Ankazobe
FMNH	165492	-18.1683	47.2767	KX015600	KX015468	KX015644	KX015512	KX015556	Ambohitantely	Réserve Spéciale d'Ambohitantely, 24 km NE Ankazobe
FMNH	165493	-18.1683	47.2767	KX015603	KX015471	KX015647	KX015515	KX015559	Ambohitantely	Réserve Spéciale d'Ambohitantely, 24 km NE Ankazobe
FMNH	165565	-18.1683	47.2767	KX015602	KX015470	KX015646	KX015514	KX015558	Ambohitantely	Réserve Spéciale d'Ambohitantely, 24 km NE Ankazobe
FMNH	165566	-18.1067	47.2517	KX015601	KX015469	KX015645	KX015513	KX015557	Ambohitantely	NW of Réserve Spéciale d'Ambohitantely, 28 km NNE Ankazobe
FMNH	166157	-19.7067	47.8350	KX015595	KX015463	KX015639	KX015507	KX015551	Ankilahila	Forêt de Ankilahila, along Andrindrimbolo River, 16.2 km SE Tsinjoarivo
FMNH	166158	-19.7067	47.8350	KX015593	KX015461	KX015637	KX015505	KX015549	Ankilahila	Forêt de Ankilahila, along Andrindrimbolo River, 16.2 km SE Tsinjoarivo
FMNH	170794	-21.2900	47.4333	KX015594	KX015462	KX015638	KX015506	KX015550	Ranomafana	Parc National de Ranomafana, Vatoharanana, 4 km SW Ranomafana (ville)
O. hova, south clade										
FMNH	156601	-24.5617	46.7217	KX015605	KX015473	KX015649	KX015517	KX015561	Andohahela	Réserve Naturelle Intégrale d'Andohahela, parcel 1, 20 km SE Andranondambo
FMNH	162052	-22.4833	46.9683	KX015606	KX015474	KX015650	KX015518	KX015562	Ivohibe	Réserve Spéciale d'Ivohibe, 8 km E Ivohibe
FMNH	162053	-22.4217	46.8983	KX015608	KX015476	KX015652	KX015520	KX015564	Ivohibe	Réserve Spéciale d'Ivohibe, 8 km NE Ivohibe, 5.5 km SE Angodongodona
FMNH	162054	-22.4217	46.8983	KX015607	KX015475	KX015651	KX015519	KX015563	Ivohibe	Réserve Spéciale d'Ivohibe, 8 km NE Ivohibe, 5.5 km SE Angodongodona
FMNH	162058	-22.4267	46.9383	KX015609	KX015477	KX015653	KX015521	KX015565	Ivohibe	Réserve Spéciale d'Ivohibe, 9 km NE Ivohibe, 6.5 km ESE Angodongodona
FMNH	165705	-22.1712	46.9459	KX015610	KX015478	KX015654	KX015522	KX015566	Andringitra	Parc National d'Andringitra, 8.5 km SE Antanifotsy
FMNH	165984	-22.1712	46.9459	KX015611	KX015479	KX015655	KX015523	KX015567	Andringitra	Parc National d'Andringitra, 8.5 km SE Antanifotsy
FMNH	167972	-22.2117	46.8450	KX015612	KX015480	KX015656	KX015524	KX015568	Andringitra	Parc National d'Andringitra, Forêt de Ravaro, 12.5 km SW Antanifotsy
FMNH	167973	-22.2117	46.8450	KX015613	KX015481	KX015657	KX015525	KX015569	Andringitra	Parc National d'Andringitra, Forêt de Ravaro, 12.5 km SW Antanifotsy
FMNH	167974	-22.2117	46.8450	KX015615	KX015483	KX015659	KX015527	KX015571	Andringitra	Parc National d'Andringitra, Forêt de Ravaro, 12.5 km SW Antanifotsy
FMNH	167975	-22.2117	46.8450	KX015616	KX015484	KX015660	KX015528	KX015572	Andringitra	Parc National d'Andringitra, Forêt de Ravaro, 12.5 km SW Antanifotsy
FMNH	167976	-22.2117	46.8450	KX015618	KX015486	KX015662	KX015530	KX015574	Andringitra	Parc National d'Andringitra, Forêt de Ravaro, 12.5 km SW Antanifotsy

Supplementary Table 2.1 (continued).

Collection	Catalog #	Latitude	Longitude	APP	BRCA1	ND2	RAG2	vWF	Locality	Locality Description
FMNH	167977	-22.2117	46.8450	KX015614	KX015482	KX015658	KX015526	KX015570	Andringitra	Parc National d'Andringitra, Forêt de Ravaro, 12.5 km SW Antanifotsy
FMNH	167978	-22.2117	46.8450	KX015617	KX015485	KX015661	KX015529	KX015573	Andringitra	Parc National d'Andringitra, Forêt de Ravaro, 12.5 km SW Antanifotsy
FMNH	169239	-21.8350	46.9633	KX015619	KX015487	KX015663	KX015531	KX015575	Mandriandry	Mandriandry, 4.4 km SW Tolongoina
FMNH	169240	-21.8350	46.9633	KX015620	KX015488	KX015664	KX015532	KX015576	Mandriandry	Mandriandry, 4.4 km SW Tolongoina
FMNH	178673	-23.8383	46.9583	KX015621	KX015489	KX015665	KX015533	KX015577	Midongy-Sud	Parc National de Midongy-Sud, NE slope Mont Papango, 3.5 km SW Befotaka
FMNH	178761	-23.8383	46.9583	KX015622	KX015490	KX015666	KX015534	KX015578	Midongy-Sud	Parc National de Midongy-Sud, NE slope Mont Papango, 3.5 km SW Befotaka
FMNH	178762	-23.5100	47.0517	KX015623	KX015491	KX015667	KX015535	KX015579	Midongy-Sud	Western slope Mt. Ambatobe, 1.2 km ENE Ampatramary, 9.5 km NE Midongy-Sud

Supplementary Table 2.2. Primer pairs used in this study.

Gene	Primer	Sequence (5'-3')
APP (forward)	APP-F1	TTGAGCAGATGCAGAACTAG
APP (reverse)	APP-R1	TCGCACGTTCACATGAAGCA
BRCA (forward)	BRCA-F1	GACTGAATGTAGAAAAGGCTG
BRCA (reverse)	BRCA-R2	GCTTTCTTGATAAAGTCCTCAGG
ND2 (forward)	Met-1	CTAATAAAGCTTTCGGGCCCATAC
ND2 (reverse)	ND2-LOR2	GAGTAGGCTATGATTTTDCGTA
RAG2 (forward)	Rag2-F1	ACACCAAACAATGAGCTTTC
RAG2 (reverse)	Rag2-R1	CACTGGAGACAGAGATTCCT
vWF (forward)	vWF-F1	TGGCCCGGAACTTGGYCCGCTA
vWF (reverse)	vWF-D	CCCACTCCAATGGGCACCA

Supplementary Table 2.3. Descriptions of all morphological measurements. Measurements were taken from the left side only (or right side only when the left side was damaged) using electronic calipers and a foot pedal. All craniomandibular measurements were taken as minimum distance between landmarks. Only adults defined by the presence of fully erupted permanent dentition were included. Dental loci are identified by tooth type (I/i, upper/lower incisor; C/c, upper/lower canine; P/p, upper/lower premolar; and M/m, upper/lower molar) and position, e.g. I3 refers to the phylogenetic third upper incisor of the permanent dentition.

Abbreviation	Module	Definition
CPM	Cranium	Condylopremaxillary length: Posterior-most (caudal) surface of occipital condyle to anterior-most (rostral) surface of the premaxilla.
CN	Cranium	Condylonasal length: Caudal surface of occipital condyle to anterodorsal-most surface of the nasal.
LCP	Cranium	Lambdoid crest extension: Greatest distance between caudal surface of lambdoid crest and rostral surface of the premaxilla.
CIL	Cranium	Condyloincisive length: Posterior-most (caudal) surface of occipital condyle to anterior-most (rostral) surface of I1.
CI2	Cranium	Condylo-I2 length: Posterior-most (caudal) surface of occipital condyle to anterior-most (rostral) surface of I2.
CC1	Cranium	Condylo-C1 length: Caudal surface of occipital condyle to rostral surface of C1
CP2	Cranium	Condylo-P2 length: Caudal surface of occipital condyle to rostral surface of P2.
CP3	Cranium	Condylo-P3 length: Posterior-most (caudal) surface of occipital condyle to anterior-most (rostral) surface of P3.
I2W	Cranium	Greatest width across I2.
C1W	Cranium	Greatest width across C1.
P2W	Cranium	Greatest width across P2.
P3W	Cranium	Greatest width across P3.
P4WP	Cranium	Greatest posterior width across P4: Greatest breadth across P4 as measured from labial surface of distostyle.
M1WP	Cranium	Greatest posterior width across M1: Greatest breadth across M1 as measured from lateral surface of distostyle.
M2WA	Cranium	Greatest anterior width across M2: Greatest breadth across M1 as measured from lateral surface of anterior ectostyle and/or mesiostyle.
M2WP	Cranium	Greatest posterior width across M2: Greatest breadth across M2 as measured from labial surface of distostyle.
M3W	Cranium	Greatest breadth across M3: Greatest breadth across M3, as measured from labial surface of distostyle.
NAW	Cranium	Nasal width: Greatest breadth of anterodorsal processes of nasals.
ZN	Cranium	Zygonasal length: Caudal surface of zygomatic process of maxillary to anterodorsal-most surface of the nasal.
PZ	Cranium	Premaxillary to zygomatic length: Rostral surface of premaxilla to caudal surface of zygomatic process of maxilla.
UTR	Cranium	Upper toothrow length: Rostral surface of I1 to caudal surface of M3.
BB	Cranium	Braincase breadth: Greatest cranial breadth, as measured across squamosals.
ZB	Cranium	Zygomatic breadth: Greatest breadth across maxillary zygomatic processes.
MH	Mandible	Height of mandible: Greatest distance between coronoid and angular processes of mandible.
AC	Mandible	Angular process depth: Greatest distance between angular and condyloid processes of mandible.
CCor	Mandible	Height of coronoid process: Greatest distance between ventral surface of condyloid process and dorsal surface of coronoid process.
MCW	Mandible	Mandibular condyle width: Greatest breadth across buccal and labial surfaces of mandibular condyle.
ML	Mandible	Condylo-i1 length: Greatest distance between caudal surface of condyloid process of mandible and rostral surface of i1. ML2 of Olson et al. (2004).
Ai1	Mandible	Greatest length of mandible: Greatest distance between caudal surface of angular process of mandible and rostral surface of i1.
m3c1	Mandible	Greatest distance between m3 and c1 as measured from the rostral surface of c1 to caudal surface of m3.
m3i1	Mandible	Greatest distance between m3 and i1 as measured from the rostral surface of c1 to caudal surface of i1.

Supplementary Table 2.4. All morphological measurements used in this study. Species, sex, latitude, longitude, and locality descriptions are from museum catalog pages. Additional tabs contain the final available-case and imputed datasets which were used in analyses. Measurement abbreviations are defined in Table A3. Collection abbreviations are as follows: Field Museum of Natural History (FMNH); Université d'Antananarivo, Mention Zoologie et Biodiversité Animale (formerly Département de Biologie Animale) (UADBA); The Natural History Museum (formerly the British Museum of Natural History) (BMNH); American Museum of Natural History (AMNH); National Museum of Natural History, Smithsonian Institution (USNM); Museum of Comparative Zoology, Harvard University (MCZ); and Muséum national d'Histoire naturelle (MNHN). This table is available upon request, and will be made publicly available upon this chapter's acceptance in a peer-reviewed journal.

Supplementary Table 2.5. Variable loadings on the first two principal components using the available-case morphometric dataset. Variable loadings from the PCA conducted with the imputed dataset are provided in Table 2.4.

	PC1	PC2
AC	0.1558149	0.1505319
BB	0.1977655	0.19726346
CIL	0.2294578	-0.1646029
ML	0.2356553	-0.0950955
CC1	0.2317646	-0.1353099
CI2	0.2314973	-0.1518044
CP2	0.2277608	-0.144017
CPM	0.2223506	-0.1936897
MH	0.2147953	0.06642977
LCP	0.1530232	-0.3298899
M1WP	0.193934	0.26466588
M2WA	0.1954702	0.25313733
M2WP	0.1869189	0.28407154
M3W	0.1974073	0.24227839
m3c1	0.2197031	-0.0586383
m3i1	0.2124208	-0.1319287
MC2	0.1360846	0.22453925
NC	0.2135393	-0.2190829
ZN	0.1684105	-0.2730083
NAW	0.1761106	0.25920414
P2W	0.1897596	0.12033134
P3W	0.1398505	0.16255473
P4WP	0.1996031	0.24912765
PZ	0.1968678	-0.2020024
UTR	0.2224247	-0.1056668

Supplementary Table 2.6. Variable loadings on the first two canonical variates for two analyses: the first analysis in which the training dataset included only individuals with DNA sequence data, and the second analysis in which the training dataset included individuals that were assigned by geographic locality. Both analyses used the available-case morphometric dataset; variable loadings from analyses using the imputed morphometric dataset are provided in Table 2.5. Variables that are not shown were removed during the shrunken covariance procedure. In both analyses, the first two canonical variates explain 100% of the variance.

CVA loadings: available-case dataset with training data assigned by sequencing (Supplementary Fig. 2.11a)		
	CV1	**CV2**
CIL	-1625.1947	-393.49593
ML	143.64223	-1.802916
CC1	-169.11972	428.458882
CI2	1861.10818	532.969954
CP2	-126.94722	-412.46183
CPM	-721.63253	-1032.9352
LCP	-164.40989	161.415776
M1WP	-32.76167	19.290818
M2WA	-163.5872	-43.562121
M2WP	133.89406	18.099059
m3c1	-548.53267	23.888057
m3i1	353.22321	-10.892227
NC	644.68243	617.398428
ZN	-494.9676	-260.75669
PZ	429.52343	175.426521
UTR	293.23724	142.916469

CVA loadings: available-case dataset with training data assigned by geography (Supplementary Fig. 2.11b)		
	CV1	**CV2**
BB	50.371331	-47.656627
CIL	-236.40971	-619.2505
ML	-36.046951	23.047432
CC1	-129.6345	318.131769
CI2	22.269239	292.299357
CP2	178.18226	-189.60906
CPM	313.872148	238.983327
MH	-23.27074	56.717619
LCP	-125.88141	38.967278
M1WP	-10.475781	57.473184
M2WA	53.346311	-104.28223
M2WP	-48.805679	19.533284
M3W	-41.34469	9.71585
m3c1	-220.45653	-54.51055
m3i1	127.73082	-14.187977
NC	9.439722	-246.77909
ZN	-10.054008	95.907387
NAW	-26.245646	4.415501
P2W	-31.343035	-20.728895
P3W	45.288425	8.137785
P4WP	-10.618336	31.810516
PZ	48.343328	-149.74219
UTR	55.345687	180.288112

Supplementary Table 2.7. Classification matrices for two canonical variate analyses: the first analysis in which the training dataset included only individuals with DNA sequence data, and the second analysis in which the training dataset included individuals that were assigned by geographic locality. Both analyses used the available-case morphometric dataset; classification matrices from analyses using the imputed morphometric dataset are provided in Table 2.6.

Classification matrix for CVA where training data are assigned by sequencing (Supplementary Fig. 2.11a)				
	North	**Central**	**South**	**% Correct**
North	9	0	0	100%
Central	0	7	0	100%
South	0	0	13	100%
Classification matrix for CVA where training data are assigned by geography (Supplementary Fig. 2.11b)				
	North	**Central**	**South**	**% Correct**
North	12	4	0	75.00%
Central	4	25	1	83.33%
South	0	1	26	96.30%

Supplementary Table 2.8. Predicted individual clade memberships and probabilities from four canonical variate analyses: (CVA1) analyzed using the imputed dataset with sequenced individuals used as training data; (CVA2) analyzed using the available-case dataset with sequenced individuals used as training data; (CVA3) analyzed using the imputed dataset, with individuals assigned to a training dataset based on north, central, or south localities; and (CVA4) analyzed using the imputed dataset, with individuals assigned to a training dataset based on North, Central, or South localities. Individuals highlighted in green produced consistent clade assignments across all four analyses, and were used to plot projected range limits (Fig. 2.7).

						CVA trained using molecular assignments			CVA trained using geographic assignments		
Museum	**Catalog #**	**Species**	**Sex**	**Latitude**	**Longitude**	***a priori* assignment**	**CVA1 (probability)**	**CVA2 (probability)**	***a priori* assignment**	**CVA3 (probability)**	**CVA4 (probability)**
AMNH	275361	*O. hova*	female	-13.69	49.44	-	South(1.0)	-	North	North(0.943)	-
FMNH	166237	*O. hova*	male	-14.00	48.43	North	North(1.0)	North(1.0)	North	North(0.88)	North(0.981)
FMNH	166238	*O. hova*	male	-14.00	48.43	North	North(1.0)	North(1.0)	North	North(0.772)	North(0.976)
FMNH	166239	*O. hova*	female	-14.00	48.43	North	North(1.0)	North(1.0)	North	North(0.598)	North(0.613)
FMNH	166240	*O. hova*	female	-14.00	48.43	North	North(1.0)	North(1.0)	North	North(0.553)	North(0.995)
FMNH	166208	*O. hova*	male	-14.02	48.42	North	North(1.0)	North(1.0)	North	North(0.968)	North(0.995)
AMNH	275189	*O. hova*	female	-14.15	48.96	-	North(1.0)	North(1.0)	North	North(0.796)	North(0.984)
AMNH	275190	*O. hova*	female	-14.17	48.95	-	North(1.0)	-	North	North(0.699)	-
AMNH	275191	*O. hova*	female	-14.42	49.15	-	South(0.99)	Central(0.853)	North	North(0.976)	North(0.899)
UADBA	46344	*O. hova*	male	-14.42	49.15	-	Central(1.0)	Central(1.0)	North	South(0.44)	North(0.509)
FMNH	159578	*O. hova*	male	-14.44	49.74	North	North(1.0)	North(1.0)	North	North(0.608)	Central(0.534)
FMNH	159695	*O. hova*	male	-14.44	49.78	North	North(1.0)	North(1.0)	North	North(0.857)	Central(0.749)
UADBA	10731	*O. hova*	female	-14.44	49.78	-	North(1.0)	North(1.0)	North	North(0.937)	North(0.981)
BMNH	48.67	*O. hova*	male	-14.49	49.77	-	North(1.0)	North(1.0)	North	North(0.993)	North(0.996)
BMNH	48.68	*O. hova*	male	-14.49	49.77	-	North(1.0)	-	North	North(0.737)	-
BMNH	48.69	*O. hova*	female	-14.49	49.77	-	Central(1.0)	Central(1.0)	North	North(0.99)	North(0.969)
AMNH	275360	*O. hova*	female	-14.53	49.43	-	North(0.923)	-	North	North(0.989)	-
FMNH	167480	*O. hova*	male	-14.54	49.44	-	North(1.0)	-	North	Central(0.89)	-
FMNH	167481	*O. hova*	male	-14.54	49.44	North	North(1.0)	North(1.0)	North	Central(0.516)	North(0.823)
UADBA	32518	*O. hova*	male	-14.54	49.44	-	Central(1.0)	Central(0.997)	North	North(0.676)	North(0.393)
FMNH	167482	*O. hova*	female	-14.61	49.43	North	North(1.0)	North(1.0)	North	North(0.691)	North(0.735)
FMNH	156161	*O. hova*	unknown	-14.78	49.45	-	North(1.0)	North(0.802)	-	North(0.916)	North(0.986)

Supplementary Table 2.7 (continued).

Museum	Catalog #	Species	Sex	Latitude	Longitude	*a priori* assignment	CVA1 (probability)	CVA2 (probability)	*a priori* assignment	CVA3 (probability)	CVA4 (probability)
FMNH	176460	*O. hova*	male	-15.23	49.53	-	North(1.0)	North(1.0)	-	North(0.579)	Central(0.839)
FMNH	176477	*O. hova*	male	-15.23	49.53	-	North(1.0)	North(0.957)	-	North(0.832)	Central(1.0)
FMNH	176455	*O. hova*	male	-15.40	49.44	-	North(1.0)	North(1.0)	-	North(0.985)	Central(0.815)
FMNH	176456	*O. hova*	male	-15.40	49.44	-	North(1.0)	North(1.0)	-	North(0.947)	North(0.905)
BMNH	48.7	*O. hova*	male	-15.42	49.82	-	North(1.0)	-	-	North(0.63)	-
BMNH	48.71	*O. hova*	male	-15.42	49.82	-	North(1.0)	-	-	North(0.977)	-
BMNH	48.72	*O. hova*	male	-15.42	49.82	-	Central(1.0)	-	-	North(0.999)	-
BMNH	48.75	*O. hova*	female	-15.42	49.82	-	North(0.999)	-	-	North(0.959)	-
BMNH	48.76	*O. hova*	female	-15.42	49.82	-	North(1.0)	North(1.0)	-	North(0.755)	North(0.726)
BMNH	48.77	*O. hova*	female	-15.42	49.82	-	North(1.0)	-	-	North(0.951)	-
BMNH	48.64	*O. hova*	male	-15.70	49.50	-	North(1.0)	-	-	North(0.649)	-
BMNH	48.65	*O. hova*	male	-15.70	49.50	-	Central(0.999)	North(0.999)	-	North(0.932)	Central(0.97)
BMNH	48.66	*O. hova*	male	-15.70	49.50	-	Central(1.0)	-	-	Central(0.793)	-
UADBA	16203	*O. hova*	unknown	-16.31	49.77	-	Central(0.999)	South(0.992)	-	South(0.875)	North(0.968)
UADBA	16202	*O. hova*	unknown	-16.41	49.66	-	North(0.999)	North(0.957)	-	North(0.737)	South(0.851)
BMNH	1991.248	*O. hova*	male	-16.85	49.13	-	North(1.0)	North(1.0)	-	North(0.955)	North(0.973)
UADBA	48489	*O. hova*	female	-18.01	48.02	-	North(1.0)	North(0.763)	-	North(0.871)	North(0.946)
UADBA	49948	*O. hova*	male	-18.01	48.02	-	North(1.0)	North(0.994)	-	Central(0.921)	North(0.514)
UADBA	49989	*O. hova*	female	-18.01	48.02	-	North(1.0)	North(1.0)	-	North(0.522)	Central(0.993)
FMNH	165490	*O. hova*	male	-18.17	47.28	Central	Central(1.0)	Central(1.0)	Central	Central(0.989)	Central(0.849)
FMNH	165491	*O. hova*	male	-18.17	47.28	-	North(1.0)	-	Central	South(0.494)	-
FMNH	165493	*O. hova*	male	-18.17	47.28	-	Central(1.0)	-	Central	Central(0.994)	-
FMNH	165565	*O. hova*	female	-18.17	47.28	Central	Central(1.0)	Central(1.0)	Central	Central(0.98)	Central(0.996)
UADBA	10391	*O. hova*	male	-18.17	47.28	-	Central(1.0)	Central(1.0)	Central	Central(0.999)	Central(0.999)
UADBA	47870	*O. hova*	unknown	-18.31	48.02	-	North(1.0)	North(1.0)	Central	Central(0.951)	Central(0.989)
UADBA	48464	*O. hova*	male	-18.31	48.02	-	North(1.0)	-	Central	Central(0.976)	-
UADBA	48652	*O. hova*	female	-18.31	48.02	-	North(1.0)	North(1.0)	Central	North(0.686)	Central(0.95)

Supplementary Table 2.7 (continued).

Museum	Catalog #	Species	Sex	Latitude	Longitude	*a priori* assignment	CVA1 (probability)	CVA2 (probability)	*a priori* assignment	CVA3 (probability)	CVA4 (probability)
UADBA	49730	*O. hova*	male	-18.31	48.02	-	North(1.0)	North(1.0)	Central	Central(0.973)	Central(0.969)
UADBA	49739	*O. hova*	male	-18.31	48.02	-	North(1.0)	North(0.988)	Central	Central(0.553)	North(0.635)
UADBA	49981	*O. hova*	male	-18.31	48.02	-	North(1.0)	-	Central	Central(0.918)	-
UADBA	45709	*O. hova*	unknown	-18.40	47.94	-	North(1.0)	North(1.0)	Central	Central(0.921)	Central(0.99)
UADBA	45809	*O. hova*	unknown	-18.40	47.94	-	North(1.0)	North(0.967)	Central	Central(0.916)	Central(0.997)
FMNH	188699	*O. hova*	male	-18.42	47.94	-	North(1.0)	North(1.0)	Central	Central(0.958)	Central(0.995)
UADBA	33279	*O. hova*	male	-18.43	48.79	-	Central(1.0)	-	Central	Central(0.912)	-
UADBA	33306	*O. hova*	male	-18.43	48.79	-	North(1.0)	Central(0.937)	Central	Central(0.809)	Central(0.946)
UADBA	47867	*O. hova*	unknown	-18.47	47.96	-	North(1.0)	North(1.0)	Central	Central(0.685)	Central(0.935)
UADBA	32680	*O. hova*	male	-18.80	48.41	-	South(0.999)	-	Central	South(0.755)	-
UADBA	48167	*O. hova*	unknown	-18.81	48.98	-	Central(1.0)	Central(1.0)	Central	Central(0.957)	Central(0.979)
MNHN	1984.523	*O. hova*	female	-18.83	48.45	-	Central(1.0)	Central(1.0)	Central	Central(0.924)	Central(1.0)
UADBA	32568	*O. hova*	female	-18.84	48.39	-	Central(0.854)	South(0.922)	Central	Central(0.95)	Central(0.989)
UADBA	45615	*O. hova*	male	-18.84	48.35	-	Central(0.999)	Central(0.973)	Central	Central(0.864)	Central(0.992)
UADBA	11352	*O. hova*	male	-18.85	48.36	-	Central(1.0)	North(0.916)	Central	Central(0.682)	Central(0.536)
UADBA	33193	*O. hova*	female	-18.85	48.31	-	North(0.872)	Central(1.0)	Central	Central(0.602)	Central(0.535)
BMNH	48.84	*O. hova*	female	-18.93	48.42	-	North(1.0)	-	Central	Central(0.862)	-
FMNH	209142	*O. hova*	female	-18.98	48.46	-	North(1.0)	North(1.0)	Central	North(0.512)	Central(0.993)
FMNH	226060	*O. hova*	male	-19.04	48.35	-	Central(0.999)	Central(1.0)	Central	North(0.971)	North(0.936)
FMNH	166157	*O. hova*	male	-19.71	47.84	Central	Central(1.0)	Central(1.0)	Central	Central(0.995)	Central(0.993)
FMNH	166158	*O. hova*	female	-19.71	47.84	Central	Central(1.0)	Central(1.0)	Central	Central(0.645)	Central(0.679)
UADBA	11509	*O. hova*	male	-19.71	47.84	-	Central(0.914)	Central(0.981)	Central	Central(0.873)	Central(0.997)
UADBA	45007	*O. hova*	unknown	-20.19	47.70	-	South(0.925)	Central(1.0)	Central	South(0.769)	Central(1.0)
UADBA	45890	*O. hova*	unknown	-20.19	47.70	-	Central(1.0)	Central(0.998)	Central	North(0.837)	Central(0.902)
FMNH	161744	*O. hova*	male	-20.78	47.17	Central	Central(1.0)	Central(1.0)	Central	Central(0.985)	Central(0.976)
FMNH	161745	*O. hova*	female	-20.78	47.17	Central	Central(1.0)	Central(1.0)	Central	Central(0.925)	Central(0.943)
FMNH	161746	*O. hova*	male	-20.78	47.17	-	Central(0.995)	-	Central	Central(0.975)	-

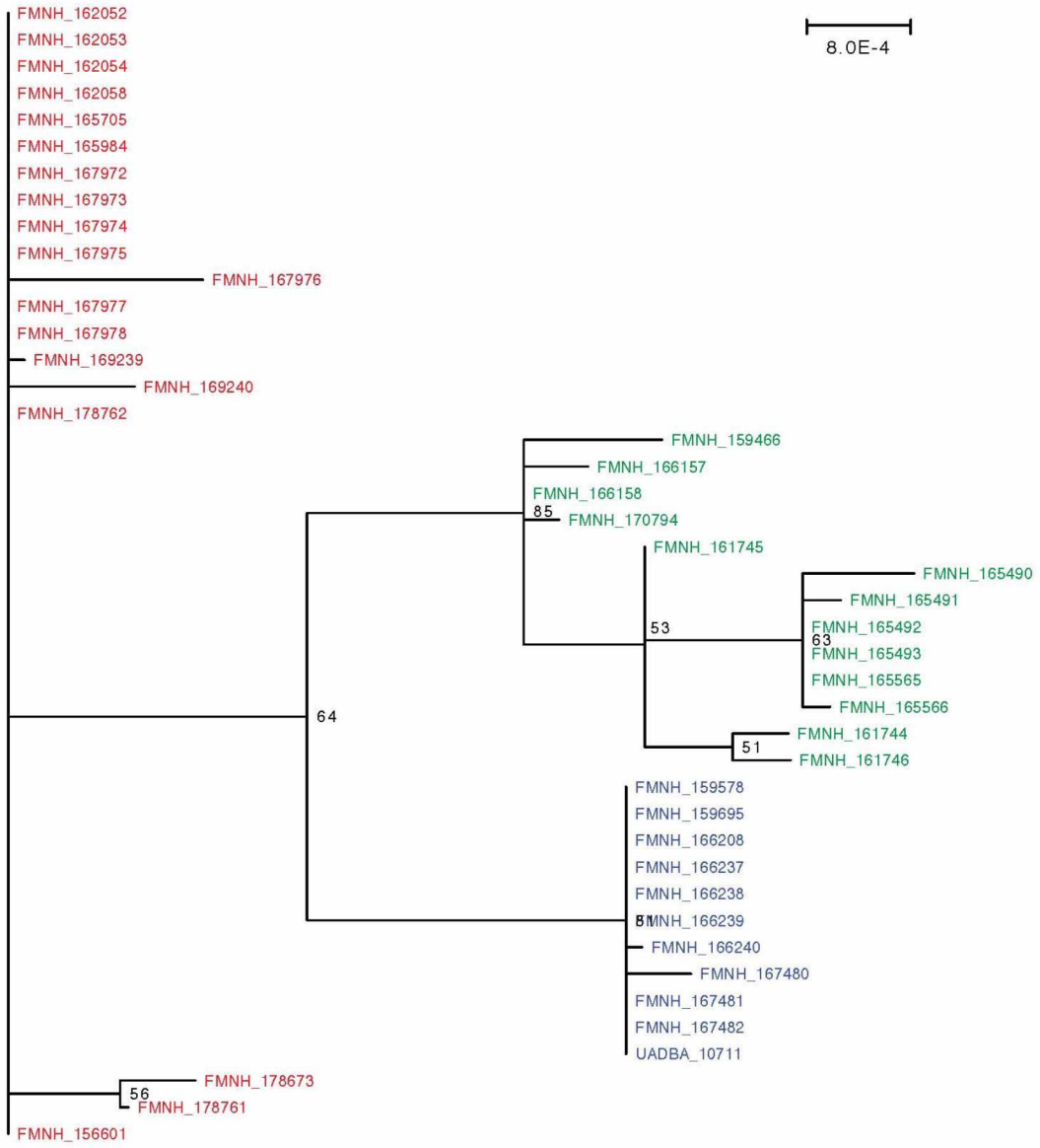
Supplementary Table 2.7 (continued).

Museum	Catalog #	Species	Sex	Latitude	Longitude	*a priori* assignment	CVA1 (probability)	CVA2 (probability)	*a priori* assignment	CVA3 (probability)	CVA4 (probability)
UADBA	30295	*O. hova*	unknown	-20.78	47.16	-	South(1.0)	South(0.921)	Central	South(0.709)	Central(0.877)
FMNH	170794	*O. hova*	male	-21.29	47.43	Central	Central(1.0)	Central(1.0)	Central	North(0.669)	Central(0.995)
UADBA	11764	*O. hova*	unknown	-21.29	47.43	-	Central(1.0)	Central(1.0)	Central	Central(0.928)	Central(0.999)
FMNH	169239	*O. hova*	male	-21.59	47.50	South	South(1.0)	South(1.0)	South	Central(0.926)	South(0.769)
UADBA	48395	*O. hova*	male	-22.19	46.97	-	Central(1.0)	South(0.988)	South	South(0.695)	South(0.572)
FMNH	167972	*O. hova*	male	-22.21	46.85	South	South(1.0)	South(1.0)	South	South(0.99)	South(1.0)
FMNH	167973	*O. hova*	male	-22.21	46.85	-	South(1.0)	-	South	South(1.0)	-
FMNH	167974	*O. hova*	male	-22.21	46.85	South	South(1.0)	South(1.0)	South	South(1.0)	South(1.0)
FMNH	167975	*O. hova*	male	-22.21	46.85	South	South(1.0)	South(1.0)	South	South(0.922)	South(0.988)
FMNH	167976	*O. hova*	male	-22.21	46.85	South	South(1.0)	South(1.0)	South	South(0.996)	South(1.0)
FMNH	167977	*O. hova*	female	-22.21	46.85	South	South(1.0)	South(1.0)	South	South(0.966)	South(1.0)
FMNH	167978	*O. hova*	male	-22.21	46.85	South	South(1.0)	South(1.0)	South	South(0.994)	South(1.0)
FMNH	161887	*O. hova*	female	-22.42	46.90	-	Central(0.999)	North(0.999)	South	South(0.715)	South(1.0)
FMNH	162053	*O. hova*	male	-22.42	46.90	South	South(1.0)	South(1.0)	South	South(0.947)	South(0.986)
FMNH	162055	*O. hova*	male	-22.42	46.90	-	South(1.0)	South(1.0)	South	South(0.996)	South(1.0)
FMNH	162056	*O. hova*	male	-22.42	46.90	-	South(1.0)	South(1.0)	South	South(0.957)	South(1.0)
FMNH	162057	*O. hova*	male	-22.42	46.90	-	South(1.0)	South(1.0)	South	South(1.0)	South(1.0)
FMNH	161888	*O. hova*	female	-22.43	46.94	-	North(1.0)	North(1.0)	South	South(0.712)	South(0.999)
FMNH	162058	*O. hova*	male	-22.43	46.94	South	South(1.0)	South(1.0)	South	South(1.0)	South(1.0)
FMNH	162052	*O. hova*	female	-22.48	46.97	South	South(1.0)	South(1.0)	South	South(0.996)	South(1.0)
BMNH	48.8	*O. hova*	male	-22.50	47.00	-	South(1.0)	-	South	South(0.929)	-
BMNH	48.81	*O. hova*	female	-22.50	47.00	-	North(1.0)	-	South	South(0.96)	-
FMNH	183964	*O. hova*	male	-23.19	47.72	-	South(1.0)	South(1.0)	South	South(0.999)	South(1.0)
FMNH	178762	*O. hova*	male	-23.51	47.05	South	South(1.0)	South(1.0)	South	South(0.988)	South(0.999)
FMNH	178673	*O. hova*	male	-23.84	46.96	South	South(1.0)	South(1.0)	South	South(0.979)	South(0.999)
FMNH	178761	*O. hova*	male	-23.84	46.96	South	South(1.0)	South(1.0)	South	South(0.996)	South(0.997)
UADBA	30234	*O. hova*	unknown	-23.84	46.96	-	South(1.0)	South(1.0)	South	South(0.995)	South(1.0)

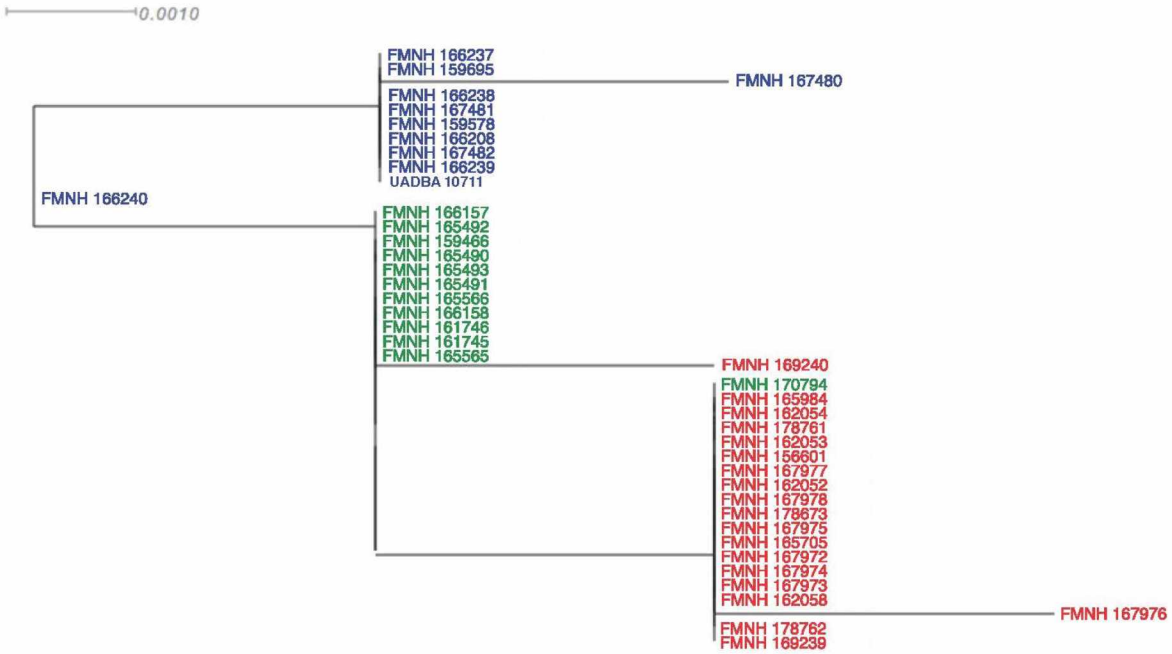
Supplementary Table 2.7 (continued).

Museum	Catalog #	Species	Sex	Latitude	Longitude	*a priori* assignment	CVA1 (probability)	CVA2 (probability)	*a priori* assignment	CVA3 (probability)	CVA4 (probability)
FMNH	156485	*O. hova*	male	-24.56	46.72	-	South(1.0)	South(1.0)	South	South(0.775)	South(0.995)
FMNH	156487	*O. hova*	male	-24.56	46.72	-	South(1.0)	South(1.0)	South	South(0.965)	South(1.0)
FMNH	156601	*O. hova*	male	-24.56	46.72	-	Central(0.975)	-	South	South(0.833)	-
FMNH	156602	*O. hova*	female	-24.56	46.72	-	North(1.0)	Central(1.0)	South	South(0.905)	South(0.991)
FMNH	156603	*O. hova*	female	-24.56	46.72	-	South(0.952)	South(0.985)	South	South(0.85)	South(1.0)
UADBA	49008	*O. hova*	female	-24.57	46.73	-	South(1.0)	South(1.0)	South	South(0.989)	South(1.0)
UADBA	10536	*O. hova*	male	-24.97	46.97	-	North(1.0)	North(1.0)	South	South(0.984)	South(0.818)
MNHN	1887874	*O. hova**	female			-	North(1.0)	North(1.0)	-	North(0.921)	Central(0.899)
MNHN	19851616	*O. hova***	female			-	North(1.0)	North(1.0)	-	North(0.671)	North(0.977)
UADBA	19130	*O. hova*	unknown			-	South(1.0)	-	-	South(0.841)	-

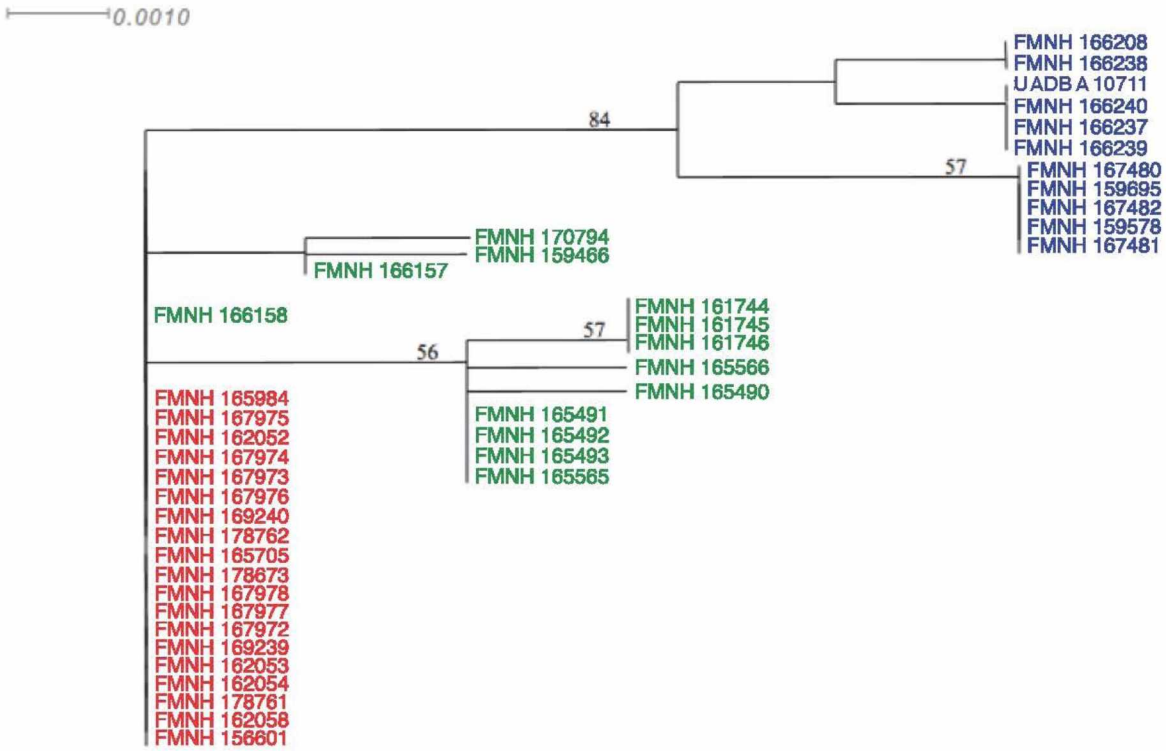
2.12 Supplementary Figures



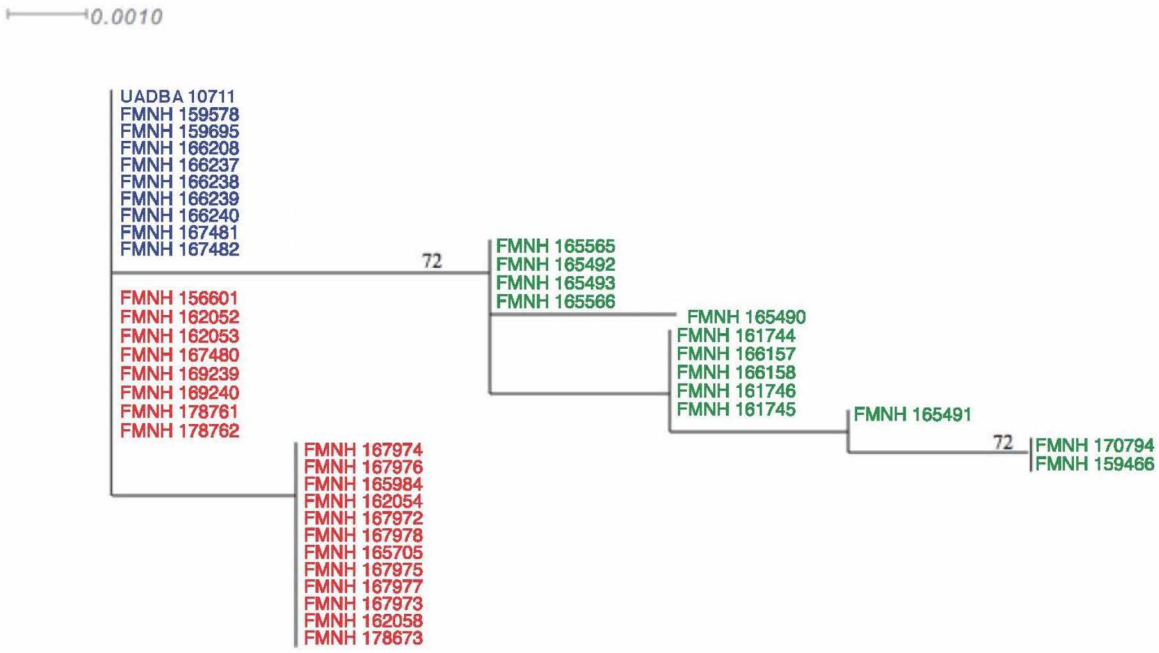
Supplementary Figure 2.1. Maximum-likelihood phylogeny from a concatenated analysis of all four nuclear loci for *O. hova*. Bootstrap values <50 are not indicated on the tree. Outgroup (*O. tetradactylus*) not shown. Specimen labels are colored according to the latitudinal clade assigned by the STRUCTURE analysis: North (blue), Central (green), and South (red).



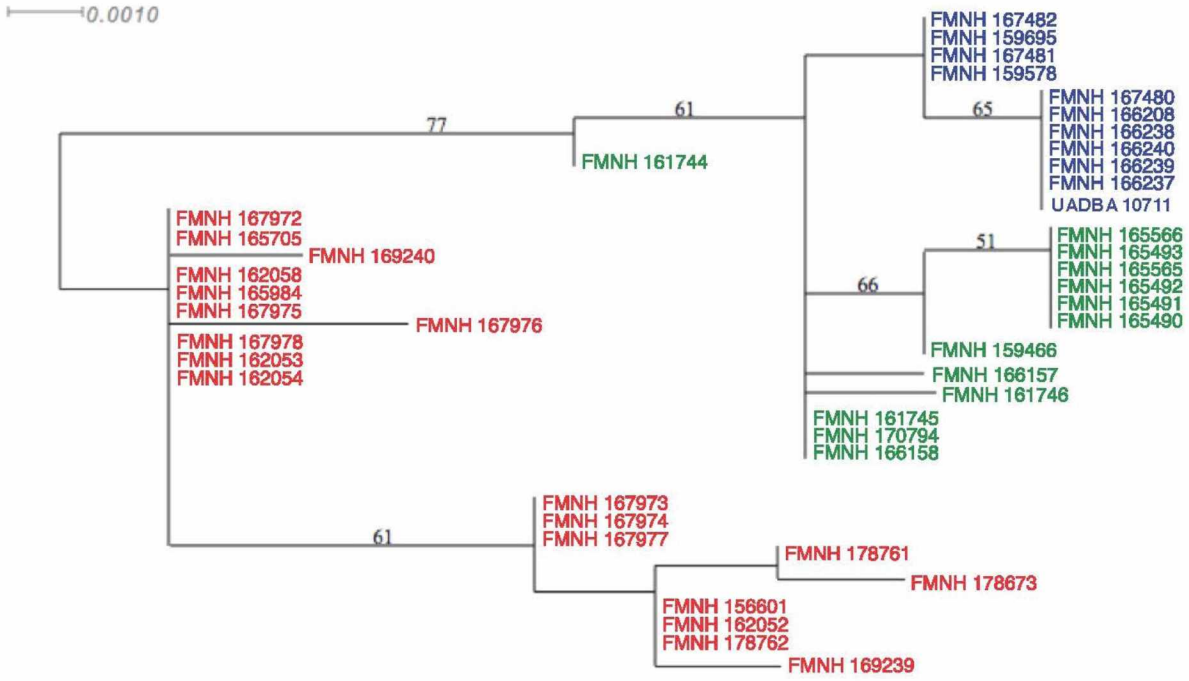
Supplementary Figure 2.2. Maximum-likelihood phylogeny of 367 bp of amyloid beta A4 precursor protein (APP) for *O. hova*. Labels and colors as in Figure A.1.



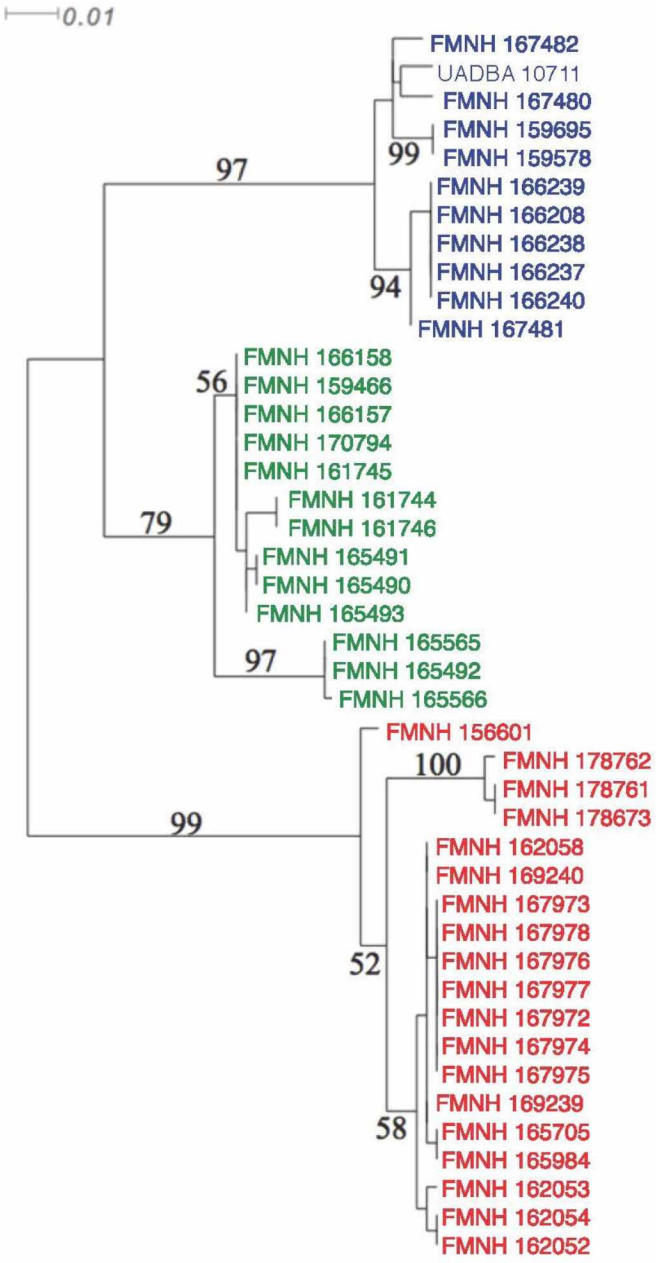
Supplementary Figure 2.3. Maximum-likelihood phylogeny of 617 bp of nuclear exon BRCA1 (similar to BRCA1 gene exon 11) for *O. hova*. Labels and colors as in Figure A.1.



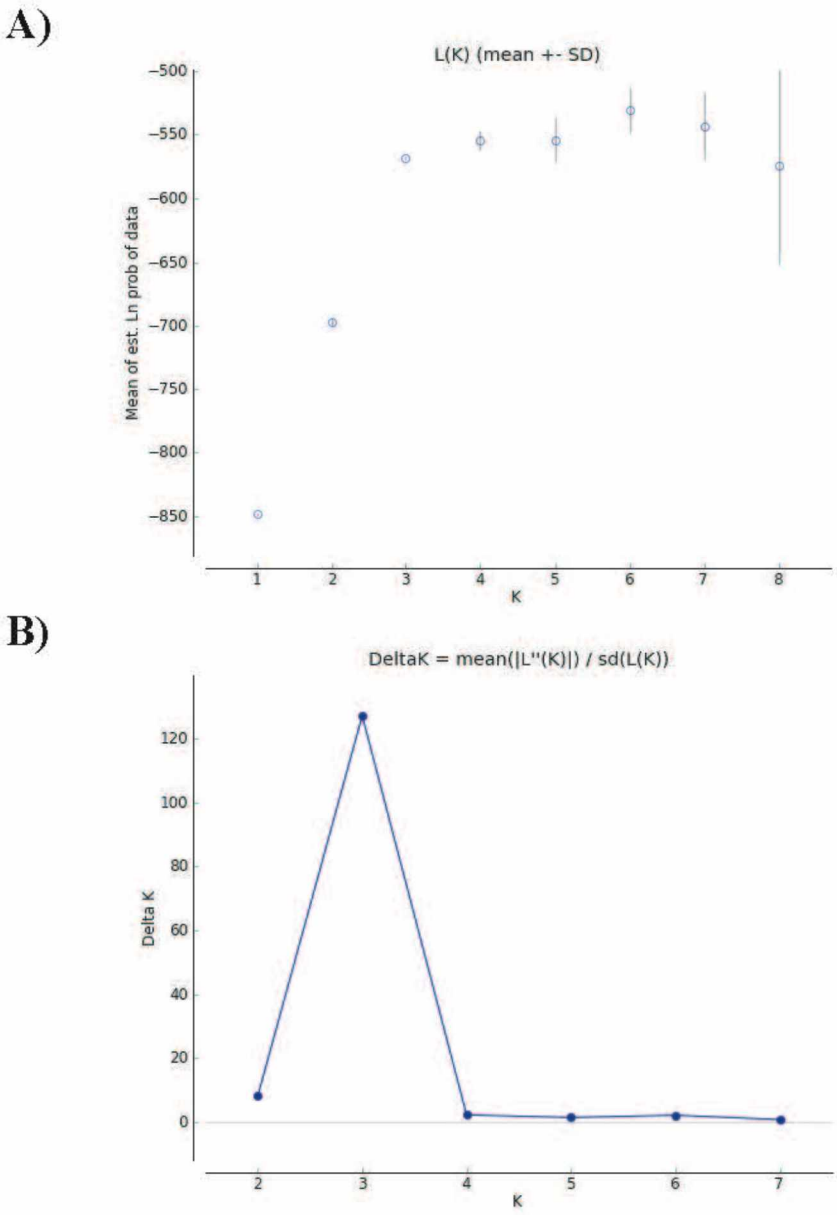
Supplementary Figure 2.4. Maximum-likelihood phylogeny of 400 bp of recombination activating gene 2 (Rag2) for *O. hova*. Labels and colors as in Figure A.1.



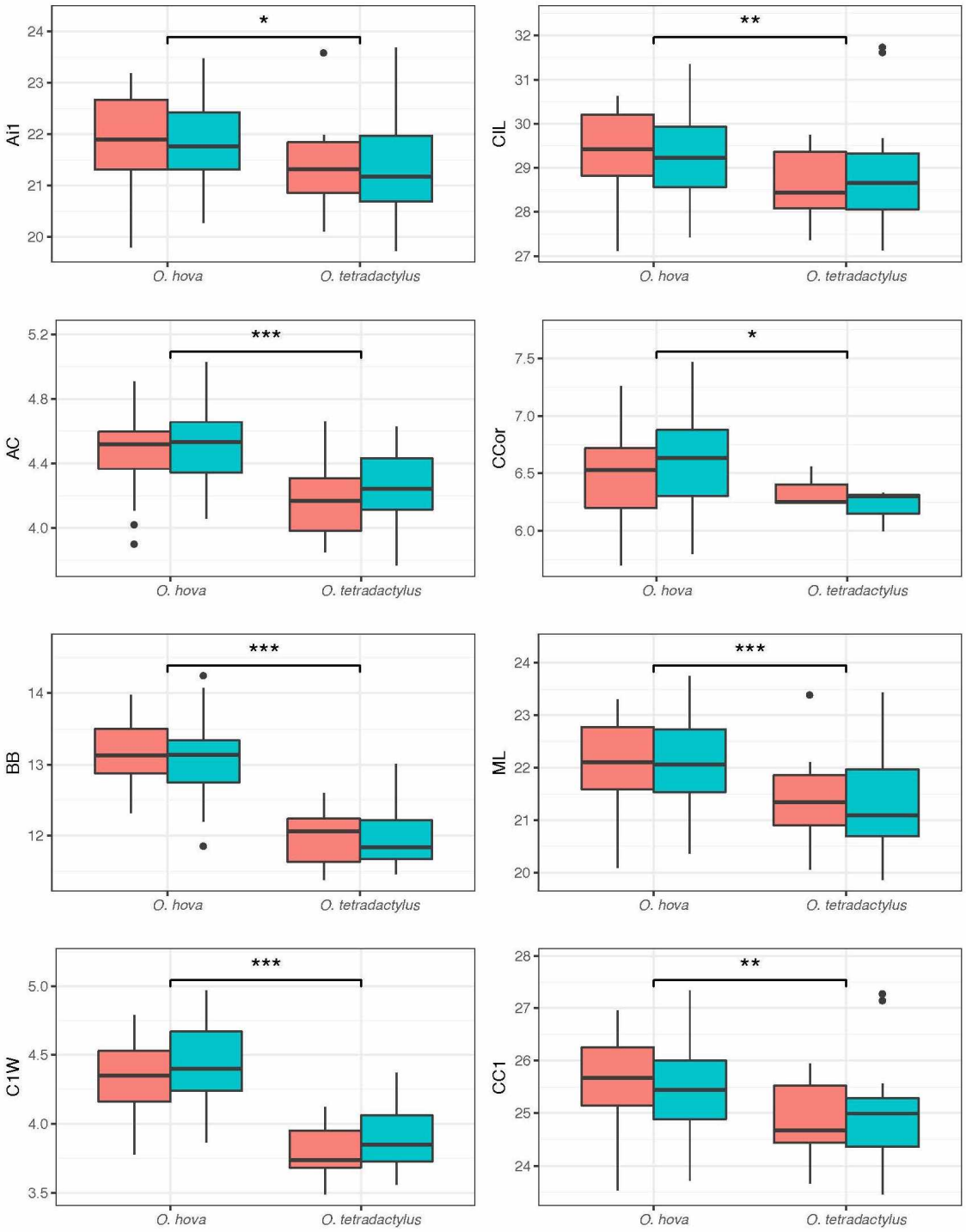
Supplementary Figure 2.5. Maximum-likelihood phylogeny of 410-530 bp of exon 28 of the von Willebrand Factor precursor (vWF) for *O. hova*. Labels and colors as in Figure A.1.



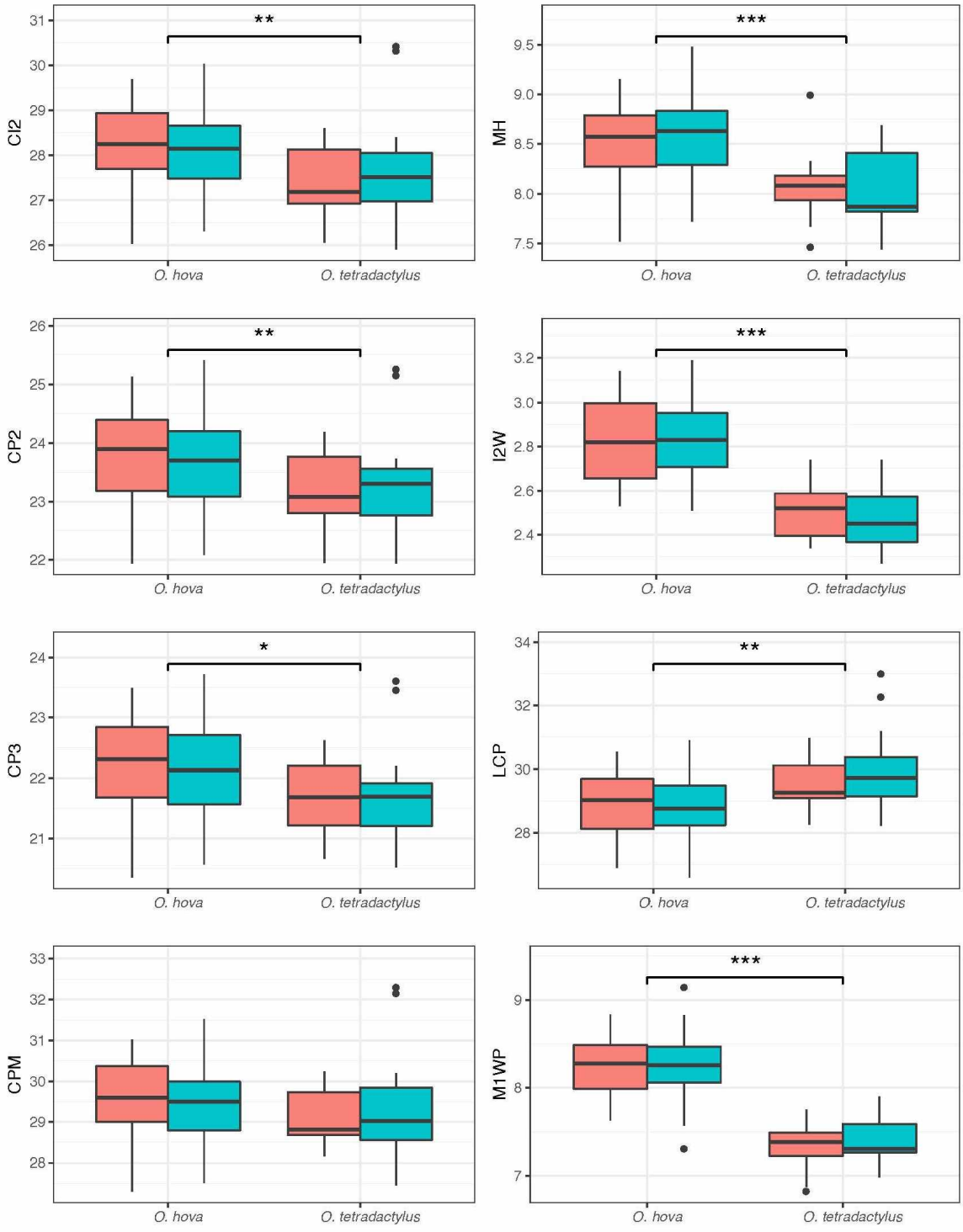
Supplementary Figure 2.6. Maximum-likelihood phylogeny of 523 bp of the mitochondrial gene NADH dehydrogenase subunit 2 (ND2) for *O. hova*. Labels and colors as in Supplementary Figure 2.1.



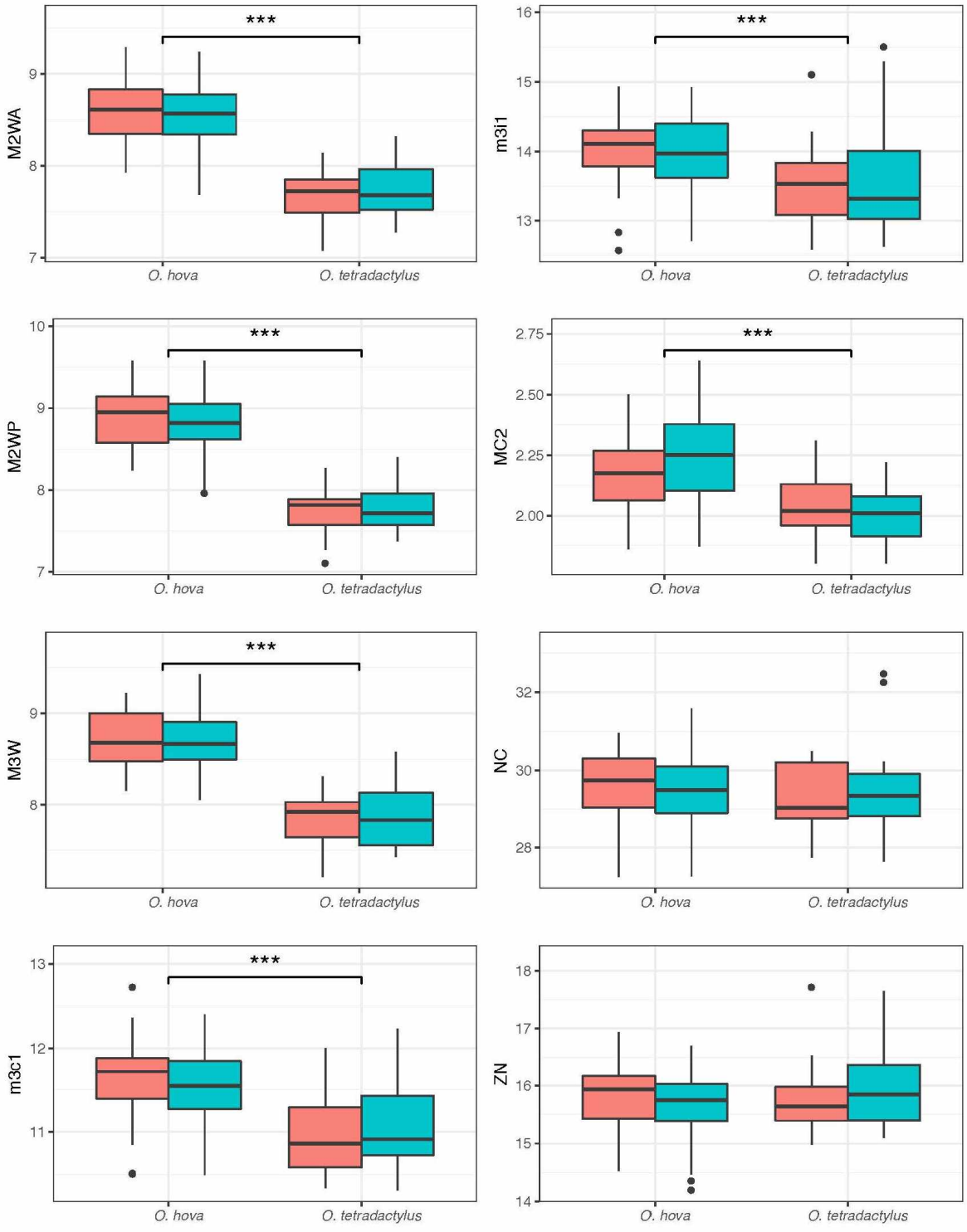
Supplementary Figure 2.7. A graphical representation of the method used for detecting the number of populations (K) in our STRUCTURE analysis (STRUCTURE result shown in Fig. 2): (A) Mean L(K) (± SD) over 10 runs for each K value, and (B) ΔK calculated as m|L''(K)|/s[L(K)]. The modal value of this distribution is the "true" K or level of structure, here 3 clusters. Figures generated by StructureHarvester (Earl and vonHoldt 2012).



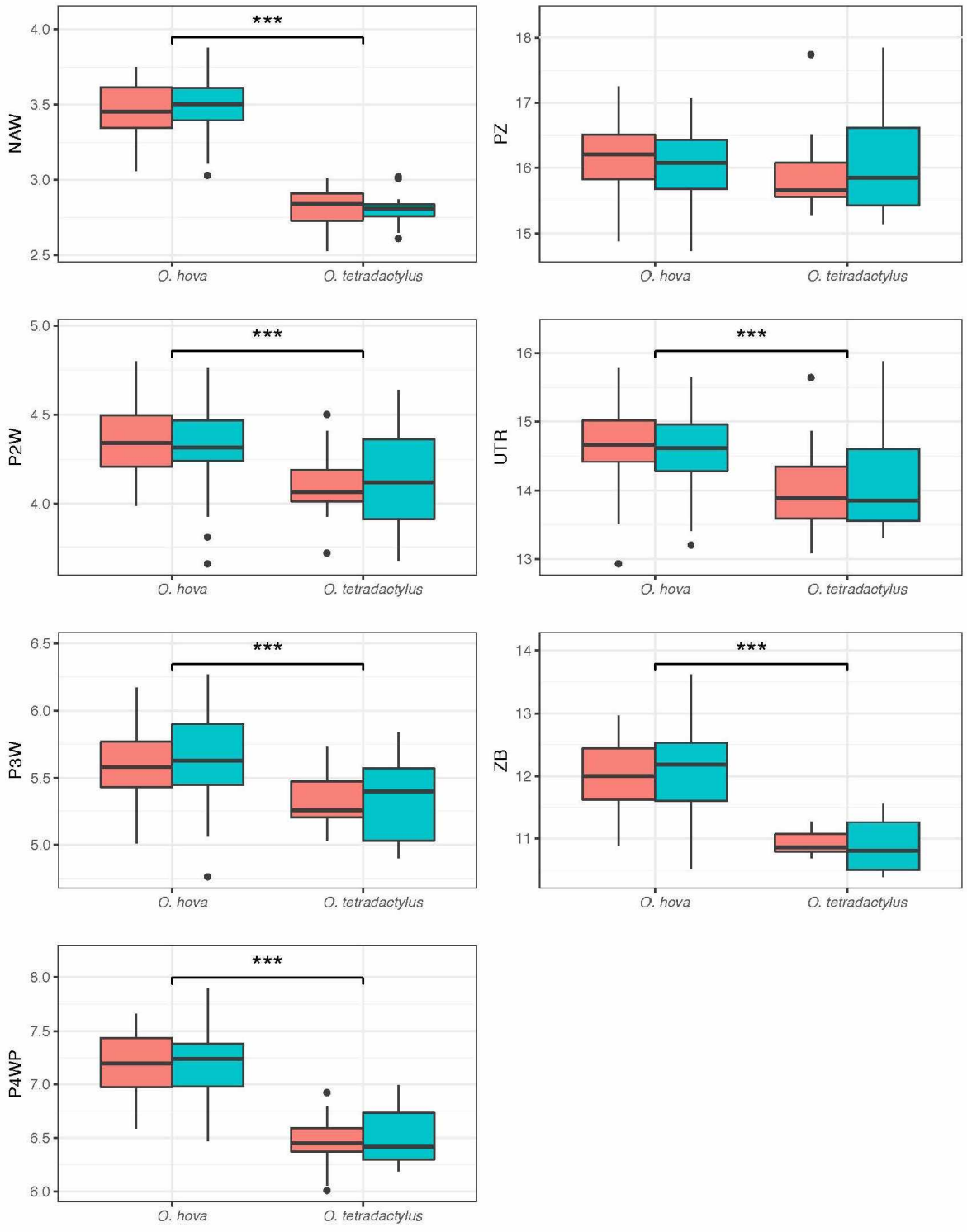
Supplementary Figure 2.8. Boxplots showing differences between *O. hova* and *O. tetradactylus*, and between males and females, for each craniomandibular measurement. Measurement abbreviations are listed on the y-axis of each plot, and descriptions of each measurement are provided in Supplementary Table 2.3. All measurement units are millimeters. Differences that are significantly different (ANOVA, alpha level = 0.05) are indicated above each bar in brackets: * = 0.05 > p > 0.01, ** = 0.01 > p > 0.001, *** = p < 0.001.



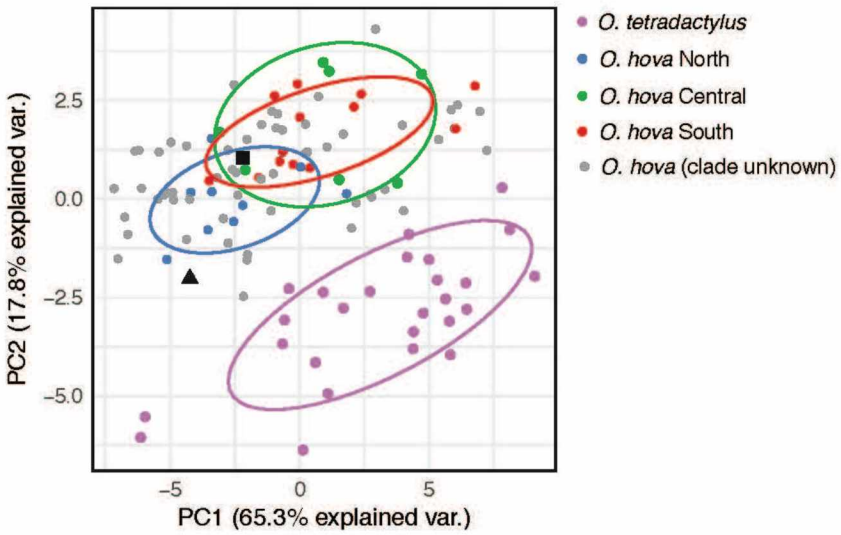
Supplementary Figure 2.8 (continued).



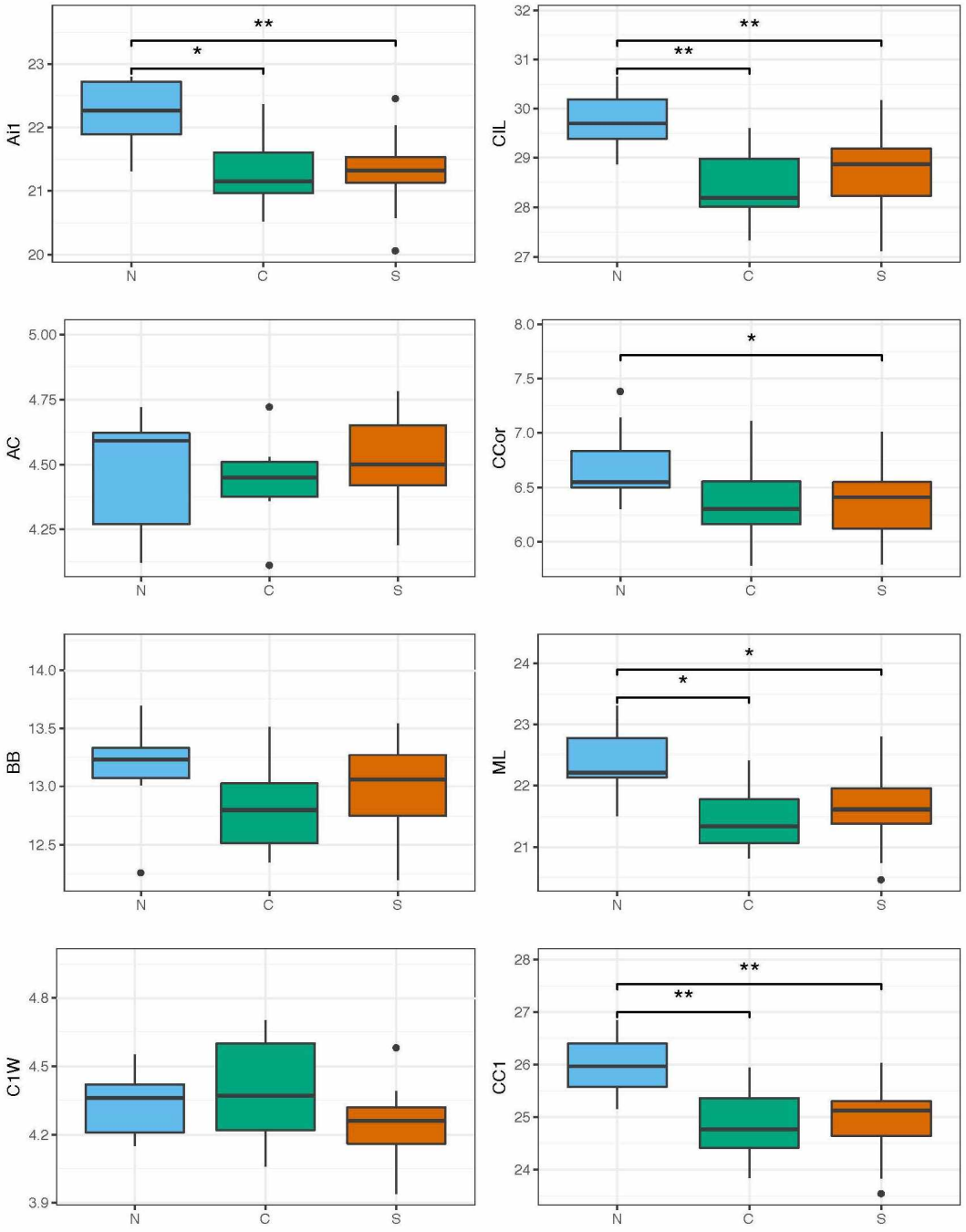
Supplementary Figure 2.8 (continued).



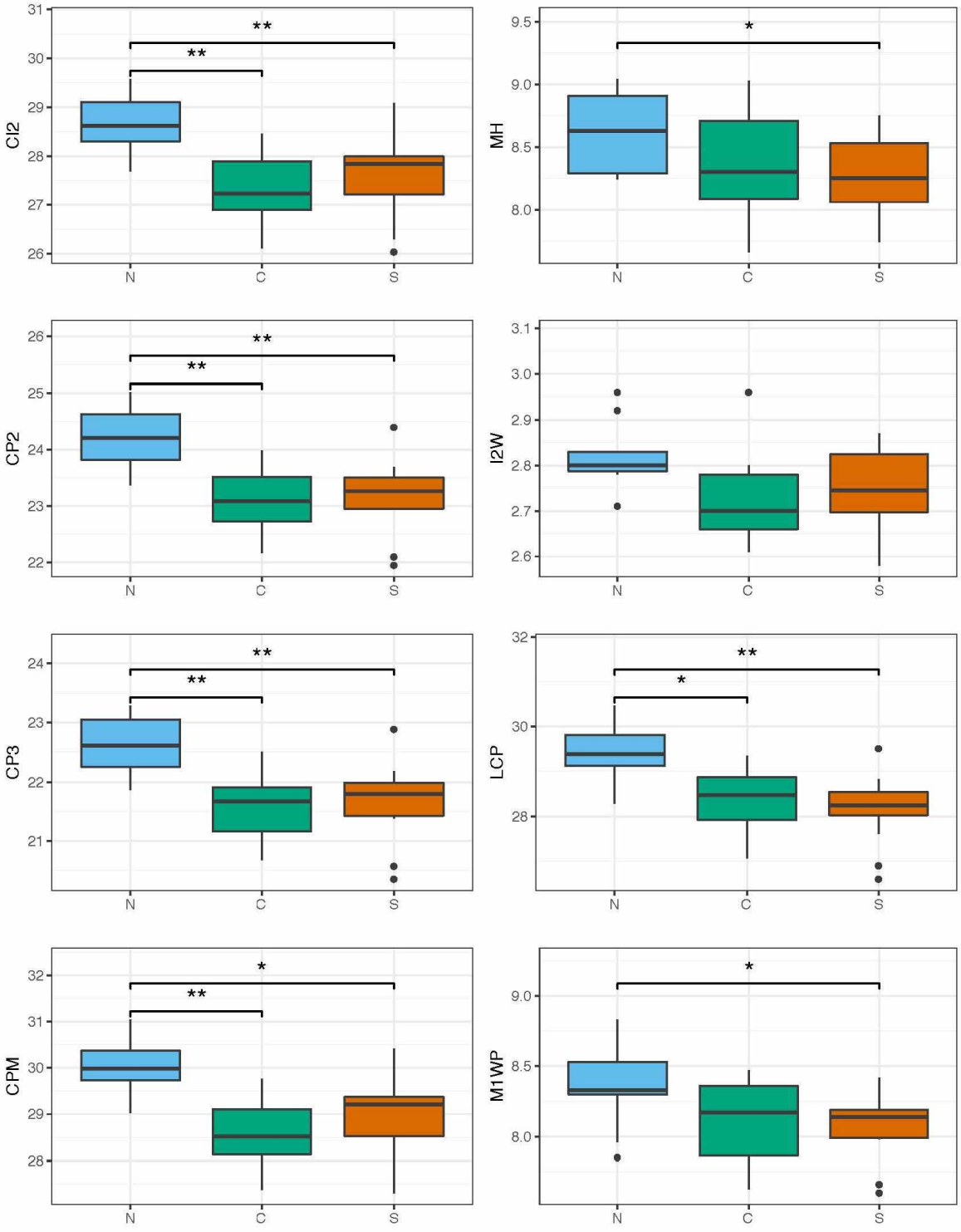
Supplementary Figure 2.8 (continued).



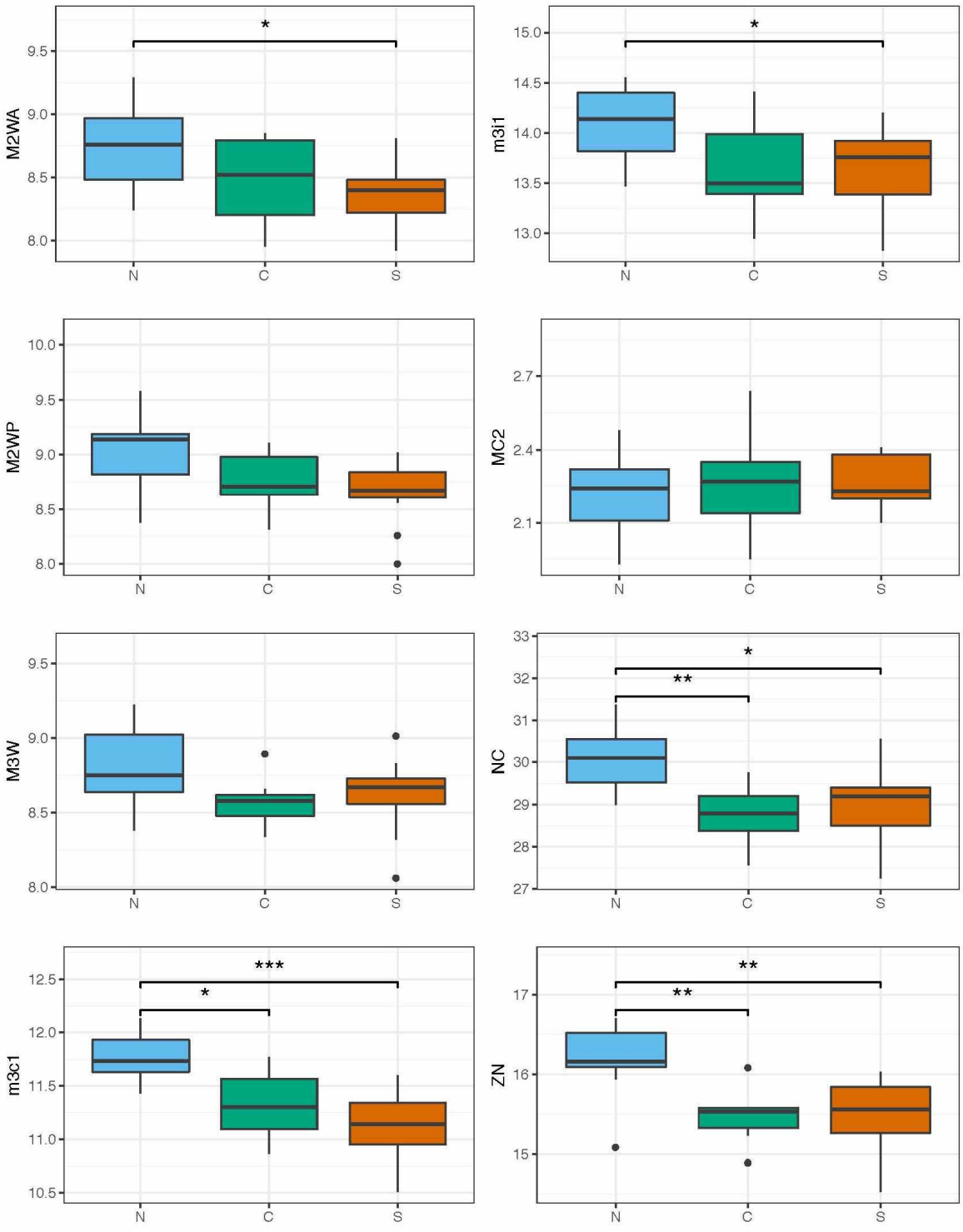
Supplementary Figure 2.9. Bivariate plots of the first two principal components using the complete, available-case craniomandibular dataset. Results from analyses using the imputed dataset are provided in Fig. 5. Individual points and 95% confidence ellipses and colored according to assignment to the North (blue), Central (green), or South (red) clade. Type specimens of *O. hova* and *O. talpoides* are indicated by a square and a triangle, respectively.



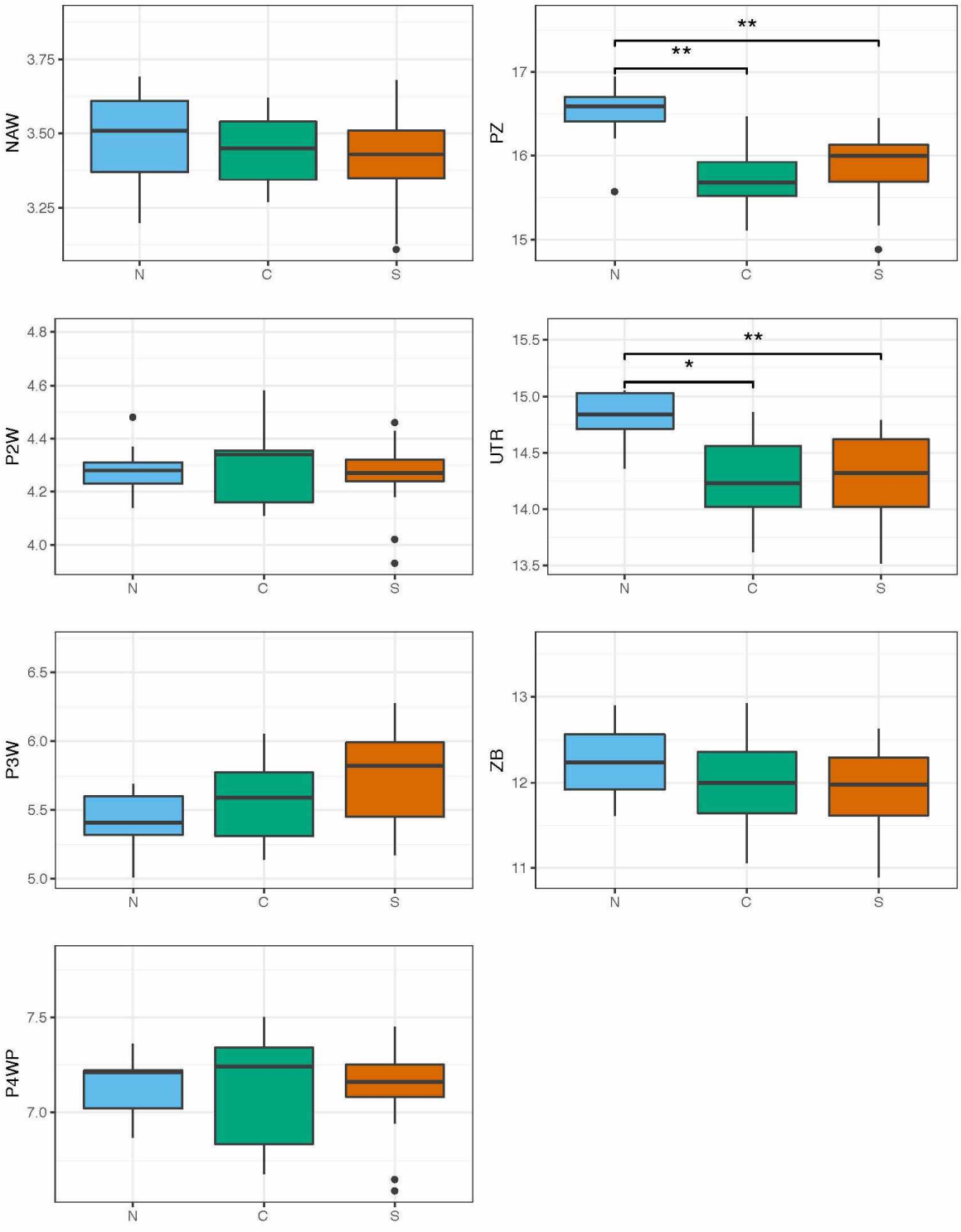
Supplementary Figure 2.10. Boxplots showing differences between the three *O. hova* clades for each craniomandibular measurement. Measurement abbreviations are listed on the y-axis of each plot, and descriptions of each measurement are provided in Supplementary Table 2.3. All measurement units are millimeters. Differences that are significantly different (ANOVA, alpha level = 0.05) are indicated above each bar in brackets: * = 0.05 > p > 0.01, ** = 0.01 > p > 0.001, *** = p < 0.001.



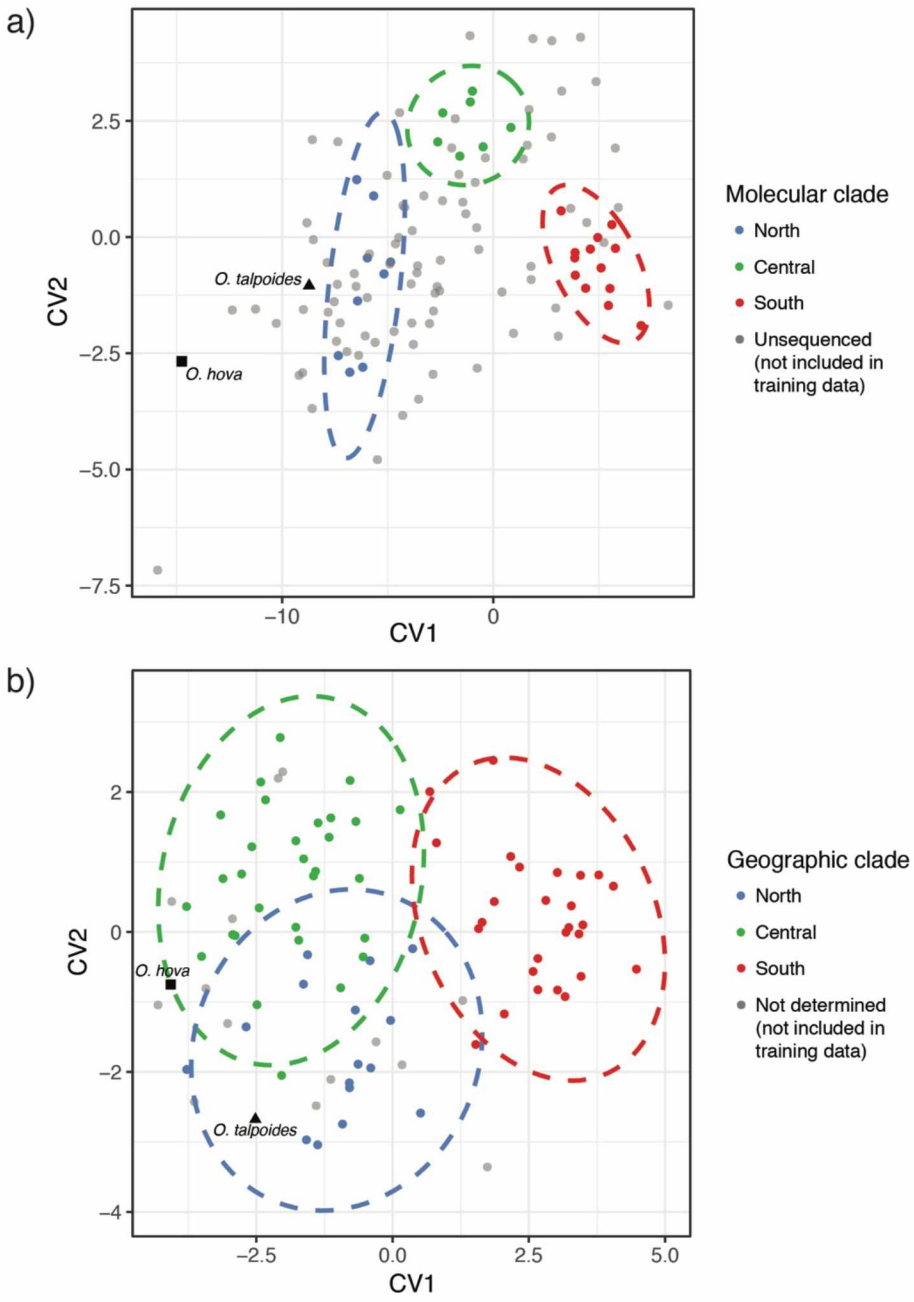
Supplementary Figure 2.10 (continued).



Supplementary Figure 2.10 (continued).



Supplementary Figure 2.10 (continued).



Supplementary Figure 2.11. Bivariate plots of the first two canonical variates where the available-cases dataset was used in the analysis in which (a) sequenced individuals were used to train the canonical variates, and (b) canonical variates were trained based on the geographic locations of individuals. Results from analyses using the imputed dataset are provided in Fig. 2.6. Individual points and 95% confidence ellipses and colored according to assignment to the North (blue), Central (green), or South (red) clade. Type specimens of *O. hova* and *O. talpoides* are indicated by a square and a triangle, respectively.

Chapter 3 – Montane regions drive diversification in Madagascar's humid-forest terrestrial vertebrates[3]

[3] Everson, K.M., Jansa, S.A., Goodman, S.M., & Olson, L.E. In preparation for *PNAS*.

3.1 Abstract

Madagascar is home to an astounding number of endemic species, many the result of asynchronous overwater dispersals followed by adaptive radiation. A growing body of evidence has suggested that mountains played an important role in *in situ* diversification by functioning as species pumps, historical refugia, and/or centers of endemism. Using a comparative phylogeographic approach, we assessed whether the breaks between four highland regions – the northern, central, and southern highlands as well as Montagne d'Ambre in the far north – correspond to genetic breaks between populations in humid-forest species. Using genetic data from 20 small mammals and five reptiles codistributed along the island's eastern humid-forest belt, we found a mountain-driven phylogeographic pattern that is remarkably congruent across taxa. This pattern is consistent regardless of the species' taxonomic group, body size, or elevational distribution. The phylogeographic divergence events we identified are also temporally concordant across taxa, suggesting that climatic conditions during the Pleistocene caused an initial isolation of the northern highlands followed by separation of the central and southern montane regions (although the same conditions do not seem to have driven the divergence of populations surrounding Mt. d'Ambre). We also identified several cases of cryptic endemism and incipient speciation, which have important conservation and evolutionary implications. This work highlights critical gaps in geographic sampling, reaffirms the urgent need for continued collecting, and illuminates protected areas of particular importance for harboring and generating diversity in one of the world's foremost and threatened biodiversity hotspots.

3.2 Introduction

Madagascar, the island-continent off the southeast coast of Africa, is a natural laboratory for evolutionary inquiry and one of the world's hottest biodiversity hotspots (1). Over its >88-million-year history of isolation, numerous taxonomic groups have colonized the island and undergone *in situ* radiations (2–4). As a result, Madagascar harbors remarkably high levels of species richness and endemism, including 100% endemism of its 101 species of native non-volant mammals, 99% of its 199 native amphibians, and 92% of its 340 native reptiles (5, 6), although these species counts are likely vastly underestimated (7, 8). Species discovery has been particularly active in recent years, largely due to increased collecting effort in new areas and the application of new molecular and morphometric techniques (9–13). This surge in biodiversity research has produced a wealth of new data, primarily mitochondrial DNA (mtDNA) sequences from hundreds of localities and taxa, providing a unique opportunity for large-scale, multi-species analyses that can illuminate broad patterns of inter- and intraspecific variation. Such research is urgently important in light of rapid and accelerating habitat loss, which has already resulted in the depletion of almost 40% of Madagascar's forest habitats (14, 15).

Approximately 90% of Madagascar's biodiversity is distributed in its humid forests, which form a north-south belt along the eastern side of the island (Fig. 3.1). These forests span three highland regions defined by elevations over 900 m (16): the northern highlands (Tsaratanana, Marojejy, and other nearby massifs; maximum elevation 2,876 m), the central highlands (Ankaratra, Andringitra, and related massifs; 2,658 m), and the southern highlands (the Anosyenne Mountains and associated southern extensions; 1,959 m). There is a shift in floral and faunal community composition between the northern and central highlands, which are separated by a low-lying valley known as the Mandritsara Window (17). Thought to have been primarily grasslands in the past, the Mandritsara Window has been invoked as a dispersal barrier for some humid-forest species (e.g., 18, 19). Similarly, the central and southern highlands are recognized as discrete biogeographic units separated by the Menaharaka Window, although it is thought that flora and fauna can cross this valley more easily than the Mandritsara Window (20). Humid forests also occur on Montagne d'Ambre (Malagasy: Ambohitra), an isolated, volcanic mountain at the island's northern terminus. While its biotic composition is similar to that of the northern highlands, Montagne d'Ambre also harbors several locally endemic species (21, 22).

Researchers of tropical regions worldwide have identified several mechanisms by which mountains can promote diversification (23–26). First, as mountains rise they often form the dividing line between bioclimatic zones (27), ultimately leading to speciation as populations adapt to different habitat types. Such is the case in Madagascar, where community composition is markedly different between the humid forests on the eastern side of the island, which occur from high elevations down to sea level, and the western dry forests, which are largely shielded from rain-bearing systems off the Indian Ocean (Fig. 3.1b). For species that live at high elevations, mountain tops also function as "sky islands," where the surrounding low-elevation habitats form systems that are analogous to oceanic islands and facilitate allopatric divergence (28). Janzen (29) argued that speciation rates are particularly high in tropical mountain ranges, as these exhibit greater thermal stratification relative to temperate mountains, providing more opportunities for adaptive and vicariant diversification. Studies on Madagascar have largely supported this prediction (e.g., 30), but have also shown that the role of mountains has changed over time. For example, during Quaternary glaciation events, when the climate was cooler and drier, Madagascar's high elevation regions likely served as areas of retreat and dispersion (31, 32).

In Madagascar's eastern humid forest, a phylogeographic pattern consisting of 2 or more latitudinally distributed (north-south) haplogroups has been recognized in many species, including reptiles (33, 34), mammals (17, 35, 36), birds (37), invertebrates (38, 39), and plants (40). In many of these studies, mountains have been invoked as a likely driver of phylogeographic structure, with the authors hypothesizing that low-lying valleys have functioned as barriers to gene flow, and that these barriers may have been more pronounced in the past. However, it is not known if these results are taxon-specific, or if

they reflect overall evolutionary processes driving diversification in the humid forests. Comparative phylogeography, an approach that leverages information from multiple co-distributed species, is a powerful tool for identifying broad-scale genetic trends across a landscape. A core objective in most comparative phylogeographic studies is to search across multiple species for congruence (spatial and/or temporal) in major intraspecific divergence events. Where a congruent phylogeographic pattern is recovered across many species, a common geological and/or historical driver of diversification can be inferred and independent lines of corroborative or refutative evidence sought (41, 42).

In this study, we use a comparative phylogeographic approach across 20 small-mammal and 5 reptile species (Table 3.1) co-distributed in Madagascar's humid forests to test the hypothesis that low-elevation areas between highland regions have served as barriers to gene flow. A shared, multi-species, phylogeographic pattern associated with the breaks between high-elevation regions would demonstrate a key role of highland regions on the recent, and presumably ongoing, *in situ* diversification of Madagascar's humid-forest vertebrate taxa.

3.3 Results and Discussion

Phylogeographic structure corresponds to high-elevation regions.

We tested our phylogeographic predictions using DNA sequence data from 823 individuals representing 20 small mammals [13 tenrecs (Tenrecidae) and 7 nesomyine rodents (Nesomyinae)] endemic to Madagascar's humid forest. Examination of phylogenies for each species suggests strong latitudinal structure (Fig. 3.2) that cannot be explained by a simple isolation-by-distance model (Supp. Fig. 3.1). To assess whether this structure corresponds to breaks among high-elevation regions, we evaluated models of gene flow and population isolation using a series of n-island model tests in which each "island" was either a single highland region, or a combination of two or more adjacent regions (Fig. 3.3). In all but one of the species we evaluated, the best-fit model considers all three highlands as distinct populations, as opposed to models in which adjacent regions are panmictic (Table 3.2). Six of the species we evaluated also include specimens collected in the Montagne d'Ambre region, and these were recovered as a distinct population in 5 of 6 cases. We found the same structure when we re-analyzed published data from five reptile species (Fig. 3.4, Table 3.2, Supp. Fig. 3.2), suggesting that this strikingly congruent, mountain-driven phylogeographic pattern is widespread rather than taxon-specific.

From our n-island model analyses we also estimated rates of migration, recovering the lowest migration rates between the northern and central highlands (Fig. 3.5). This is in agreement with previous studies that have found that the area between these regions serves as an important biogeographic divide (17, 19, 20). With that said, we found that the amount of migration is very low among all regions. The percent of immigrants in each population per generation (m) ranges from <0.00001% to 0.0008%, even

between the central and southern highlands, and between the northern highlands and Montagne d'Ambre. Importantly, we were not able to distinguish whether the higher proportions of migrants are the result of smaller population sizes in Montagne d'Ambre and the southern highlands, or if they are truly the result of increased migration rates.

Body size and elevation do not affect phylogeographic structure.

While we found a relationship between phylogeographic structure and highland regions in multiple species, we also found that the strength of this relationship is variable; some species contain clades that are closely or perfectly associated with individual highlands, while other species appear to be more vagile. To determine whether this variability among species is related to ecology, we calculated two metrics for strength of phylogeographic structure [the genealogical sorting index (gsi; 43) and Φ_{ST}] and regressed these values against two data types that are readily available from museum catalog information, even in poorly studied taxa: elevation and body size. We did not identify any correlation between elevation or body size and gsi or Φ_{ST} (Fig. 3.6). Species of all body sizes, whether they occur at sea level or are restricted to high elevations, show phylogeographic structure associated with highland breaks. This result strongly suggests a historical, climatic component to this phylogeographic structure; i.e., although species with broad elevational ranges are presumably able to move among highland regions today, populations have likely been isolated in the past.

Highland phylogeographic divergence events are temporally congruent.

Using an approximate Bayesian computation (ABC) approach (44), we found that all of the species in this study share a single, synchronous divergence time between their northern- and central-highland populations, and a second, synchronous divergence between their central- and southern-highland populations (Table 3.3). These two divergence events are temporally displaced: northern- and central-highland populations diverged with the onset of the Pleistocene (mode: 2.56 MYA, 95% CI: 1.54 – 3.46 MYA), while central- and southern-highland populations diverged more recently (mode: 0.58 MYA, 95% CI: 0 – 1.20 MYA; Fig. 3.7). Neither event was caused by geological uplift, as Madagascar's highland regions had already reached their current elevations by 10-15 million years ago (45). Rather, synchronous divergence appears to have been driven by paleoclimatic changes over the last three million years. Quaternary glacial cycling (2.588 MYA - Present) is thought to have caused recurrent bouts of forest expansion and contraction into highland regions, which would have caused populations to become repeatedly isolated (31, 32). Recent evidence from environmental niche modeling has confirmed that cooler humid-forest habitats (like those in the highlands today) were much more extensive during glacial maxima, but were discontinuous during interglacial periods (46). The northern and central highlands,

which are separated by ~100 kilometers of elevations <900 m and interspersed with arid grasslands today, would have experienced the longest periods of isolation – in many taxa, we found evidence that populations never rejoined. We are not sure why the divergence time between the central and southern highlands is more recent, but it may be related to the more southern latitude, as climatic fluctuations are expected to be more pronounced in areas further from the equator (32). If populations separated and rejoined frequently across the central and southern highlands during the Quaternary, the more recent climatic cycles could have blurred or erased the first, earlier genetic signatures of divergence.

Unlike the three primary highland regions, we did not find a synchronous divergence time between populations from Montagne d'Ambre and the northern highlands across multiple species (Table 3.3). Importantly, geological evidence has shown that Montagne d'Ambre underwent a period of high volcanic activity during the Quaternary until ~2 MYA, with periodic eruptions including either cinder or lava flows (47, 48). If these caused localized extirpation, it follows that we would see stochastic, temporally staggered recolonization and subsequent divergence in different taxa.

Madagascar's humid-forest biodiversity is underestimated

Many of the species included in this study may in fact represent species complexes. We have already established the presence of undescribed species in three of the mammal taxa we evaluated in previous studies: the northern population of *Eliurus tanala* has been recognized as *E. ellermani* (Jansa et al. *under review*), the northern population of *Microgale longicaudata* has been tentatively assigned to *M prolixacaudata* pending molecular confirmation (49), and each of the three clades of *Oryzorictes hova* have been designated candidate species pending formal descriptions and name assignments (50). Other researchers have drawn similar conclusions about the reptiles we evaluated (Fig. 3.4), and have recommended future taxonomic work. Using the general mixed Yule coalescent (GMYC) model, a method that has been applied to other large-scale species delimitation studies on Madagascar (10, 51, 52), we identified 83 candidate species that warrant additional research, some of which are redundant with the candidate species that have already been identified (overview in Figs. 3.2 and 3.4, individual assignments in Supp. Tables 3.1 and 3.2). Although the GMYC has been shown to overestimate true species diversity (53), these results provide a convenient starting point for future taxonomic research. Candidate species should be tested using additional lines of evidence, including nuclear DNA and/or morphological and ecological data, plus sampling in the proposed contact zones between lineages. We echo other authors that there is an urgent need to document species-level diversity in Madagascar (7, 54, 55) as forest habitats continue to decline while species diversity remains vastly underestimated.

3.4 Conclusions

In Madagascar's eastern humid forest, we identified a phylogeographic pattern that is both spatially and temporally congruent across multiple species, which was likely driven primarily by paleoclimatic conditions during the Pleistocene. We also recognized a number of species with unique haplotypes in the Montagne d'Ambre region, although divergence of these haplotypes from their northern sister populations does not seem to have been driven by the same climatic and/or geological forces that drove diversification in the northern, central, and southern highlands.

This research lays important groundwork for future conservation efforts. Most notably, it highlights the importance of maintaining protected areas in all three highland areas and in the area around Montagne d'Ambre, as each region harbors uniqe mitochondrial haplotypes and likely as-yet undiscovered or undescribed species. In many of the species we evaluated, we also found that the regions near the Parc National d'Andringitra (22.2°S, 46.9°E) and the Parc National de Ranomafana (21.3°S, 47.4°E) harbored genetic diversity from both the central and southern highlands, thereby acting as contact zones between haplotype clades. Continued protection of these areas may therefore be inordinately beneficial for maintaining haplotype diversity, and these regions might also be suitable for studies of hybridization.

Our plots of genetic-by-geographic distance show distinct clusters of data points rather than a smooth cline (Supp. Fig. 3.1), which is a signal of population structure rather than isolation-by-distance. However, this pattern can also be caused by sampling gaps, which may have been the case for some of the taxa we evaluated. In tenrecs, for example, very few individuals have been collected from the zone between the northern and central highlands (16–18°S latitude), which could artificially amplify a signal of population structure between these regions. We recommend prioritization of biological collecting in the forested areas between the northern and central highlands to better understand which taxa are migrating and/or hybridizing between them.

We supplemented our genetic dataset with published data from reptiles, as we expected that reptiles could be easily compared to small mammals due to their similar body sizes and terrestrial lifestyles; however, a tantalizing area of future research might include additional species with diverse body sizes and ecomorphologies. Previous research has already demonstrated latitudinal phylogeographic patterns in birds (37), insects (38, 39), and even plants (40), but it will be important to demonstrate whether these patterns are congruent with those of small-bodied terrestrial animals identified here, or if they reveal a different response to low-elevation geographic breaks. We also restricted our taxonomic sampling to species with widespread distributions in the eastern humid forest, but many taxa on Madagascar are microendemic, occurring in just one or two high-elevation regions (e.g., *Brachytarsomys betsileoensis*, a nesomyine rodent that occurs only in the central highlands, or *Zonosaurus anelanelany*, a

lizard that occurs only in the southern highlands). Our work focused on understanding the forces driving population structure in broadly distributed taxa, but future research could include microendemic species to shed light on broader patterns of species richness and endemism.

Finally, this study highlights the importance of continuing biological surveys on Madagascar. Without regular inventories, large-scale geographic patterns like those identified in this study would go unnoticed, as would cryptic diversity. Conversely, the ranges of some species would be underestimated (and almost certainly are, in many cases), confounding conservation prioritization. Finally, our results further highlight the importance of biological repositories, both the specimen repositories (biological collections) that make such work repeatable, and data repositories (e.g., GenBank) that allow researchers to supplement their own datasets to identify broad and important patterns of biodiversity.

3.5 Materials and Methods

Sampling and molecular data collection

We collected mtDNA sequence data from 823 individuals representing 20 small mammals [13 tenrecs (Tenrecidae) and 7 nesomyine rodents (Nesomyinae)] endemic to Madagascar's eastern humid forest and occurring broadly throughout this region. All specimens were collected between 1987 and 2015 from 34 localities and were deposited in natural history museums (Supp. Table 3.1). Specimens were collected in strict accordance with the terms of research permits issued by the authorities in Madagascar (Ministère des Forêts and Madagascar National Parks).

For all tenrecs, 1044 bp of the mitochondrial gene NADH dehydrogenase subunit 2 (ND2) were sequenced by K.M.E., L.E.O., or other staff at the University of Alaska Museum, following previously published protocols (56). For all nesomyines, 1201 bp of the mitochondrial gene cytochrome-b (cyt-b) were sequenced by S.A.J. or other staff at the University of Minnesota following established protocols (57). All museum collections, catalog numbers, localities, and bGMYC candidate species assignments (see methods below) for tenrecs and nesomyines are provided in Supp. Table 3.1.

We supplemented these data with published mtDNA from 5 reptile species, selected using the following criteria:

1. All sequences were used in peer-reviewed publications and publicly available through GenBank.

2. All species are broadly distributed throughout Madagascar's humid forest. This was determined by visual inspection of the species' published range map from the International Union for the Conservation of Nature (IUCN, www.iucnredlist.org).

3. All DNA sequence data are associated with georeferenced voucher specimens.

4. For each species, there are a sufficient number of individuals and collection localities. We included only species with representatives from at least 10 different collection localities distributed across at least three high-elevation regions.

All GenBank accession numbers, localities, and bGMYC candidate species assignments (see methods below) for these specimens are provided in Supp. Table 3.2.

Phylogenetic and population genetic analyses

We identified the best-fit model of nucleotide substitution for each gene using ModelTest v.3.7 (58) under the Akaike Information Criterion (AIC). We then applied the top-ranking models in separate phylogenetic analyses for each species via Bayesian Inference (BI) in BEAST v.1.8.3 (59). We allowed independent evolution of branch lengths under the uncorrelated relaxed lognormal clock model and the Yule process tree prior. In each analysis we used two independent MCMC runs (4 chains each) of 10,000,000 iterations each, sampling trees every 1,000 iterations to yield a total of 10,000 trees. Runs were combined in LogCombiner, the first 20% of these trees were removed as burn-in, and the last 80,000 trees were used to construct a majority-rule consensus tree and assign Bayesian Posterior Probabilities (BPPs) to each branch using TreeAnnotator v1.6.1 (59). Phylogenies were plotted alongside locality maps using the command "phylo.to.map" in the R package phytools (60).

Within each of the species we evaluated, we also conducted a discriminant analysis of principal components (DAPC) in the R package adegenet to find and assign individuals to genetic clusters. This method has been shown to be effective for single-locus data (61). We first used the function find.clusters, which reduces the genetic data to principal components and identifies the optimal number of genetic groups using a Bayesian information criterion, then used the function dapc to assign individuals to these groups via discriminant analysis. The identified genetic clusters are indicated on Fig. 3.2.

The spatial-genetic patterns generated by population structure can be easy to mistake for isolation-by-distance (IBD , 62); therefore, we plotted genetic distance by geographic distance using the command kde2d in the R package MASS (63) to visually inspect whether genetic diversity formed a continuous cline (as expected with IBD) or if it formed distinct patches (as expected with multiple discrete populations).

Island-model analyses

We used the BI implementation of the coalescence-based program Migrate-N v. 3.6.11 (64) to test for high-elevation-driven population structure and to estimate migration rates among populations. For the purposes of this analysis, we grouped individuals according to their high-elevation region, determined using visual inspection in QGIS (65). We then tested four models of population structure for each species:

(1) three distinct populations delimited by the northern, central, and southern highlands, (2) a distinct northern population but panmictic central and southern highlands, (3) a distinct southern population but panmictic northern and central highlands, and (4) panmixia across all three highland regions (Fig. 3.3). For all analyses we sampled 10,0000 generations with an increment of 20 and a burn-in of 20,000, and we inspected effective sample size values to ensure that our sampling was sufficient. We used the Metropolis-Hastings sampling algorithm for parameter values and the default heating strategy for thermodynamic integration. We compared and ranked models using Bayes factors, which were calculated using Bezier-approximated log marginal likelihood values.

For all species that included individuals from the far northern Montagne D'Ambre region, we conducted one additional round of Migrate-N model tests in which these localities were either separate from, or panmictic with, the northern highland population; individuals from Montagne d'Ambre were not included in the first round of analyses. Along with our model test results, Migrate-N also estimated mutation-scaled migration rates among populations ($M = m/\mu$, where m is the immigration rate per generation).

Linear regressions

We used the phylogenies of each species to calculate their genealogical sorting indices (GSI; 40), which is a measure of the degree of exclusive ancestry on a scale from 0 to 1, where 1 indicates complete genetic exclusivity (i.e., monophyly). These values were calculated for each species and each highland region using the command "gsi" in the R package genealogicalSorting (66), with p-values calculated from 10,000 permutations using the command "permutationTest". We also estimated the population fixation index Φ_{ST} between adjacent highland regions using sequence data from each species in GenAlEx (67). Both GSI and Φ_{ST} values were normalized using a logit transformation, then we conducted linear regressions to determine whether either metric was significantly correlated with (a) the lowest elevation that the species is known to occur (in meters), as listed in (16) or in the species' IUCN species accounts, or (b) body size in millimeters, given as head-body length for mammals (68) or snout-vent length for reptiles (69). Body size data were log-transformed due to right-skewness. These values are provided in Supp. Table 3.3.

Phylogeographic concordance analysis

We tested for simultaneous lineage divergence under an approximate Bayesian computation (ABC) analytical framework using msBayes (44). This program is used to evaluate the number of temporally distinct divergence events that may have occurred between pairs of lineages for any number of species. We conducted separate analyses for the divergence of northern and central populations, the

central-southern divergence, and the northern-Montagne d'Ambre divergence, for all species where these populations were recovered in Migrate-N. Because our dataset includes diverse taxonomic groups and several different mitochondrial loci, we calculated relative mutation rates for each lineage/locus combination using published divergence time data, and included these relative rates as mutation scalars in msBayes (Supp. Table 3.3).

For all analyses, we first conducted a pilot msBayes analysis of 100,000 simulations with default parameter values, which we used as a guide to set realistic priors for the final analyses. The final analyses used a maximum average divergence time (τ) of 0.5 coalescent units and included 2.5 million simulations. We used categorical regression to retain the 1,000 simulations with the best fit to the empirical data, using π (nucleotide diversity), θ_w (Watterson's theta), π_{net} (Nei and Li's net nucleotide divergence), and the denominator of Tajima's *D* as summary statistics. We could not reject the scenario of simultaneous divergence if the 95% confidence interval (CI) of Ω [var(τ)/E(τ)] encompassed zero. Where we could not reject the scenario of simultaneous divergence, we calculated divergence time using the posterior distribution of E(τ) and relative mutation rates (Supp. Table 3.3).

Identification of candidate species

Candidate species were identified using bGMYC (70), an R-based Bayesian implementation of the general mixed Yule coalescent model (71). As input for these analyses, we used the posterior distributions of ultrametric trees estimated in our BEAST phylogenetic analyses, thinned at regular intervals to 100 total post-burn-in trees. Tips were pruned using the function drop.tip in the R package Phytools (60) so that only unique haplotypes were included. We ran bGMYC for 10,000 generations, sampling every 100 generations and discarding the first 1000 as burn-in. We used a conservative threshold for species delimitation (i.e., erring on the side of "lumping" rather than "splitting"), considering groups as separate candidate species only if the posterior probability that individuals are conspecific was less than 0.01.

3.6 Acknowledgements

This research was supported by grant DEB-1120904 (L.E.O.) and a Graduate Research Fellowship (K.M.E.) from the National Science Foundation; the University of Alaska Museum; a Bass Senior Fellowship awarded by the Field Museum of Natural History (L.E.O.); a Graduate Student Research Award from the Society of Systematic Biologists (K.M.E.); a Grant-in-Aid of Research from the American Society of Mammalogists (K.M.E.); and a Dissertation Completion Fellowship from the University of Alaska Fairbanks Graduate School (K.M.E). We are grateful for the specimen and tissue loans from the following individuals and institutions: Christopher J. Raxworthy, Nancy B. Simmons, and

Neil P. Duncan (American Museum of Natural History); Bruce D. Patterson, Lawrence R. Heaney, Adam W. Ferguson, and the late William T. Stanley (Field Museum of Natural History); Paula D. Jenkins, Louise Tomsett, and Richard Sabin (The Natural History Museum, London [formerly the British Museum of Natural History]); Daniel Rakotondravony, Steven M. Goodman, and Voahangy Soarimalala (Université d'Antananarivo, Département de Biologie Animale); Helen F. James, Darrin P. Lunde, and Michael D. Carleton (National Museum of Natural History, Smithsonian); Priscilla K. Tucker, Cody W. Thompson, and Philip Myers (University of Michigan Museum of Zoology). Finally, we thank the following individuals for their support and assistance during various stages of this project: Derek S. Sikes, Naoki Takebayashi, Sharon A. Jansa, Robert P. Anderson, Kyndall B.P. Hildebrandt, Krystal Fales, Savannah Anchinges, and our friend William T. Stanley.

3.7 References

1. Vences M, Wollenberg KC, Vieites DR, Lees DC (2009) Madagascar as a model region of species diversification. *Trends Ecol Evol* 24:456–465.
2. Yoder AD, Nowak MD (2006) Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. *Annu Rev Ecol Evol Syst* 37:405–431.
3. Crottini A, et al. (2012) Vertebrate time-tree elucidates the biogeographic pattern of a major biotic change around the K–T boundary in Madagascar. *Proc Natl Acad Sci USA* 109:5358–5363.
4. Samonds KE, et al. (2012) Spatial and temporal arrival patterns of Madagascar's vertebrate fauna explained by distance, ocean currents, and ancestor type. *Proc Natl Acad Sci USA* 109:5352–5357.
5. Goodman SM, Benstead JP (2005) Updated estimates of biotic diversity and endemism for Madagascar. *Oryx* 39:73–77.
6. Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GA, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403:853–858.
7. Vieites DR, et al. (2009) Vast underestimation of Madagascar's biodiversity evidenced by an integrative amphibian inventory. *Proc Natl Acad Sci USA* 106:8267–8272.
8. Perl RGB, et al. (2014) DNA barcoding Madagascar's amphibian fauna. *Amphibia-Reptilia* 35:197–206.
9. Raxworthy C, Ingram C, Rabibisoa N, Pearson RG (2007) Applications of ecological niche modeling for species delimitation: A review and empirical evaluation using day geckos (*Phelsuma*) from Madagascar. *Syst Biol* 56:907–923.
10. Monaghan MT, et al. (2009) Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Syst Biol* 58:298–311.

11. Raxworthy CJ, et al. (2003) Predicting distributions of known and unknown reptile species in Madagascar. *Nature* 426:837–841.
12. Hotaling S, et al. (2016) Species discovery and validation in a cryptic radiation of endangered primates: Coalescent-based species delimitation in Madagascar's mouse lemurs. *Mol Ecol* 25:2029–2045.
13. Olivieria G, et al. (2007) The ever-increasing diversity in mouse lemurs: Three new species in north and northwestern Madagascar. *Mol Phylogenet Evol* 43:309–327.
14. Harper G, Steininger M, Tucker C, Juhn D, Hawkins F (2007) Fifty years of deforestation and forest fragmentation in Madagascar. *Environ Conserv* 34:325–333.
15. Grinand C, et al. (2013) Estimating deforestation in tropical humid and dry forests in Madagascar from 2000 to 2010 using multi-date Landsat satellite images and the random forests classifier. *Remote Sens Environ* 139:68–80.
16. Goodman S, Raherilalao M (2014) *Atlas of Selected Land Vertebrates of Madagascar* (Association Vahatra, Antananarivo, Madagascar).
17. Raxworthy CJ, Nussbaum RA (1996) Patterns of endemism for terrestrial vertebrates in eastern Madagascar. *Biogéographie de Madagascar*, ed Lourenço W (ORSTOM, Paris), pp 369–383.
18. Goodman SM, Rakotondravony D, Randriamanantsoa HN, Rakotomalala-Razanahoera M (2009) A new species of rodent from the montane forest of central eastern Madagascar (Muridae: Nesomyinae: *Voalavo*). *Proc Biol Soc Washingt* 118:863–873.
19. Gehring P-S, Glaw F, Gehara M, Ratsoavina FM, Vences M (2013) Northern origin and diversification in the central lowlands? – Complex phylogeography and taxonomy of widespread day geckos (*Phelsuma*) from Madagascar. *Org Divers Evol* 13:605–620.
20. Guillaumet J, Betsch J, Callmander M (2008) Renaud Paulian et le programme du CNRS sur les hautes montagnes à Madagascar: étage vs domaine. *Zoosystema* 30:723–748.
21. Louis EE, et al. (2008) Revision of the mouse lemurs, *Microcebus* (Primates, Lemuriformes), of northern and northwestern Madagascar with descriptions of two new species at Montagne d'Ambre National Park and Antafondro Classified Forest. *Primate Conserv* 23:19–38.
22. Raxworthy CJ, Nussbaum RA (1994) A rainforest survey of amphibians, reptiles and small mammals at Montagne d'Ambre, Madagascar. *Biol Conserv* 69:65–73.
23. Smith BT, et al. (2014) The drivers of tropical speciation. *Nature* 515:406–409.
24. Fjeldså J, Bowie RCK, Rahbek C (2012) The role of mountain ranges in the diversification of birds. *Annu Rev Ecol Evol Syst* 43:249–265.
25. Rahbek C (1997) The relationship among area, elevation, and regional species richness in neotropical birds. *Am Nat* 149:875–902.

26. Cadena CD, et al. (2012) Latitude, elevational climatic zonation and speciation in New World vertebrates. *Proc R Soc B* 279:194–201.
27. Loffler H (1984) The importance of mountains for animal distribution, species speciation, and faunistic evolution (with special attention to inland waters). *Mt Res Dev* 4:299.
28. Mccormack JE, Huang H, Knowles LL (2009) Sky islands. *Encyclopedia of Islands*, eds Gillespie, R, Clague, D (U. California Press), pp 839–843.
29. Janzen DH (1967) Why mountain passes are higher in the tropics. *Am Nat* 101:233–249.
30. Wollenberg KC, et al. (2008) Patterns of endemism and species richness in Malagasy cophyline frogs support a key role of mountainous areas for speciation. *Evolution* 62:1890–1907.
31. Wilmé L, Goodman SM, Ganzhorn JU (2006) Biogeographic evolution of Madagascar's microendemic biota. *Science* 312:1063–1065.
32. Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philos Trans R Soc Lond B Biol Sci* 359:183–195.
33. Boumans L, Vieites DR, Glaw F, Vences M (2007) Geographical patterns of deep mitochondrial differentiation in widespread Malagasy reptiles. *Mol Phylogenet Evol* 45:822–839.
34. Nagy ZT, Glaw F, Andreone F, Wink M, Vences M (2007) Species boundaries in Malagasy snakes of the genus *Madagascarophis* (Serpentes: Colubridae sensu lato) assessed by nuclear and mitochondrial markers. *Org Divers Evol* 7:241–251.
35. Yoder AD, Heckman KL (2006) Mouse lemur phylogeography revises a model of ecogeographic constraint in Madagascar. *Primate Biogeography*, eds Lehman SM, Fleagle JG (Springer, New York), pp 255–268.
36. Pastorini J, Forstner MRJ, Martin RD (2001) Phylogenetic history of sifakas (*Propithecus*: Lemuriformes) derived from mtDNA sequences. *Am J Primatol* 53:1–17.
37. Younger JL, et al. (2018) Hidden diversity of forest birds in Madagascar revealed using integrative taxonomy. *Mol Phylogenet Evol* 124:16-26.
38. Wirta H (2009) Complex phylogeographical patterns, introgression and cryptic species in a lineage of Malagasy dung beetles (Coleoptera: Scarabaeidae). *Biol J Linn Soc* 96:942–955.
39. Linares MC, Soto-Calderón ID, Lees DC, Anthony NM (2009) High mitochondrial diversity in geographically widespread butterflies of Madagascar: A test of the DNA barcoding approach. *Mol Phylogenet Evol* 50:485–495.
40. Shapcott A, et al. (2007) Can we bring Madagascar's critically endangered palms back from the brink? Genetics, ecology and conservation of the critically endangered palm *Beccariophoenix madagascariensis*. *Bot J Linn Soc* 154:589–608.

41. Avise J (2000) *Phylogeography: the history and formation of species* (Harvard University Press, Cambridge, MA).
42. Bermingham E, Moritz C (1998) Comparative phylogeography: Concepts and applications. *Mol Ecol* 7:367–369.
43. Cummings MP, Neel MC, Shaw KL (2008) A genealogical approach to quantifying lineage divergence. *Evolution* 62:2411–2422.
44. Hickerson MJ, Stahl E, Takebayashi N (2007) msBayes: Pipeline for testing comparative phylogeographic histories using hierarchical approximate Bayesian computation. *BMC Bioinformatics* 8:268.
45. de Wit MJ (2003) Madagascar: Heads it's a continent, tails it's an island. *Annu Rev Earth Planet Sci* 31:213–248.
46. Boria RA, Olson LE, Goodman SM, Anderson RP (2017) A single-algorithm ensemble approach to estimating suitability and uncertainty: cross-time projections for four Malagasy tenrecs. *Divers Distrib* 23:196–208.
47. Karche J-P (1972) Contribution à l'étude géologique de la Montagne d'Ambre et des régions voisines du nord de Madagascar. Dissertation (Univ. Besançon).
48. Brenon P (1972) The geology of Madagascar. *Biogeography and Ecology in Madagascar*, eds Battistini R, Richard-Vindard G (Springer, Dordrecht), pp 27–86.
49. Olson LE, Goodman SM, Yoder AD (2004) Illumination of cryptic species boundaries in long-tailed shrew tenrecs (Mammalia: Tenrecidae; *Microgale*), with new insights into geographic variation and distributional constraint. *Biol J Linn Soc* 83:1–22.
50. Everson KM, Hildebrandt KBP, Goodman SM, Olson LE (2018) Caught in the act: incipient speciation across a latitudinal gradient in a semifossorial mammal from Madagascar, the mole tenrec *Oryzorictes hova* (Tenrecidae). *Mol Phylogenet Evol* 126:74–84.
51. Vuataz L, Sartori M, Gattolliat J-L, Monaghan MT (2013) Endemism and diversification in freshwater insects of Madagascar revealed by coalescent and phylogenetic analysis of museum and field collections. *Mol Phylogenet Evol* 66:979–991.
52. Gehring P-S, et al. (2012) Hiding deep in the trees: Discovery of divergent mitochondrial lineages in Malagasy chameleons of the *Calumma nasutum* group. *Ecol Evol* 2:1468–1479.
53. Paz A, Crawford AJ (2012) Molecular-based rapid inventories of sympatric diversity: A comparison of DNA barcode clustering methods applied to geography-based vs clade-based sampling of amphibians. *J Biosci* 37:887–896.
54. Yoder AD, et al. (2000) Remarkable species diversity in Malagasy mouse lemurs (Primates, *Microcebus*). *Proc Natl Acad Sci USA* 97:11325–11330.

55. Nagy ZT, Sonet G, Glaw F, Vences M (2012) First large-scale DNA barcoding assessment of reptiles in the biodiversity hotspot of Madagascar, based on newly designed COI primers. *PLoS One* 7:e34506.

56. Everson KM, Soarimalala V, Goodman SM, Olson LE (2016) Multiple loci and complete taxonomic sampling resolve the phylogeny and biogeographic history of tenrecs (Mammalia: Tenrecidae) and reveal higher speciation rates in Madagascar's humid forests. *Syst Biol* 65:890–909.

57. Goodman SM, Raheriarisena M, Jansa SA (2009) A new species of *Eliurus* Milne Edwards, 1885 (Rodentia: Nesomyinae) from the Réserve Spéciale d'Ankarana, northern Madagascar. *Bonner Zool Beiträge* 56:133–149.

58. Posada D, Crandall K (1998) Modeltest: Testing the model of DNA substitution. *Bioinformatics* 14:817–818.

59. Drummond A, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214.

60. Revell LJ (2012) PHYTOOLS: An R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol* 3:217–223.

61. Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* 11:94.

62. Meirmans PG (2012) The trouble with isolation by distance. *Mol Ecol* 21:2839–2846.

63. Venables B, Ripley B (2002) *Modern Applied Statistics with S* (Springer, New York). 4th Ed.

64. Beerli P, Palczewski M (2010) Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics* 185:313–326.

65. QGIS Development Team (2012) QGIS geographic information system: Open source geospatial foundation project. Available at: http://qgis.osgeo.org.

66. Bazinet A, Neel M, Shaw K, Cummings M (2008) The genealogical sorting index: Software and web site for quantifying the exclusivity of lineages. Available at: http://www.genealogicalsorting.org/.

67. Peakall R, Smouse PE (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6, 288-295.

68. Soarimalala V, Goodman S (2011) *Les petits mammifères de Madagascar* (Association Vahatra, Antananarivo, Madagascar).

69. Glaw F, Vences M (2007) *A field guide to the amphibians and reptiles of Madagascar* (Vences & Glaw).

70. Reid NM, Carstens BC (2012) Phylogenetic estimation error can decrease the accuracy of species delimitation: A Bayesian implementation of the general mixed Yule-coalescent model. *BMC Evol Biol* 12:196.

71. Pons J, et al. (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst Biol* 55:595–609.

72. Ratsoavina F, et al. (2013) An overview of Madagascar's leaf tailed geckos (genus *Uroplatus*): Species boundaries, candidate species and review of geographical distribution based on molecular data. *Salamandra* 49:115-148.

73. Miralles A, Vences M. (2013) New metrics for comparison of taxonomies reveal striking discrepancies among species delimitation methods in *Madascincus* lizards. *PLoS One* 8:e68242.

3.8 Tables

Table 3.1. All species included in this study, including the number of individuals collected from each high-elevation region. Northern, Central, and Southern refer to Madagascar's three primary highland regions. The locus column give the abbreviated name of the sequenced mitochondrial gene: ND2 = NADH dehydrogenase subunit 2, Cyt-b = Cytochrome-b, 16S = 16S ribosomal RNA, ND1 = NADH dehydrogenase subunit 1. References are provided when genetic data were not collected by the authors of this study.

Tenrecidae (tenrecs)	**Mt. d'Ambre**	**Northern**	**Central**	**Southern**	**Locus**	**Reference**
Microgale drouhardi	2	43	9	-	ND2	This study
Microgale fotsifotsy	4	19	13	4	ND2	This study
Microgale gymnorhyncha	-	8	30	3	ND2	This study
Microgale longicaudata	5	16	23	1	ND2	This study, (46)
Microgale majori	-	3	66	8	ND2	This study, (46)
Microgale parvula	2	5	11	2	ND2	This study
Microgale principula	-	3	5	7	ND2	This study
Microgale soricoides	-	34	23	5	ND2	This study
Microgale taiva	-	5	17	-	ND2	This study
Microgale thomasi	-	1	13	2	ND2	This study
Nesogale dobsoni	-	3	17	1	ND2	This study
Nesogale talazaci	1	5	14	1	ND2	This study
Oryzorictes hova	-	11	31	4	ND2	(47)
Nesomyidae (rodents)						
Eliurus grandidieri	-	13	15	-	Cyt-b	This study
Eliurus majori	4	17	31	8	Cyt-b	This study
Eliurus minor	-	12	50	6	Cyt-b	This study
Eliurus tanala	12	26	46	21	Cyt-b	This study
Eliurus webbi	-	13	11	15	Cyt-b	This study
Gymnuromys roberti	-	3	16	2	Cyt-b	This study
Nesomys rufus	-	4	19	2	Cyt-b	This study
Gekkonidae (geckos)						
Hemidactylus mercatorius	1	9	4	9	16S	(30)
Phelsuma lineata	3	20	279	19	16S	(9, 30)
Uroplatus sikorae	4	6	35	5	ND4	(72)
Chameleonidae (chameleons)						
Calumma nasutum	-	7	28	13	ND2	(53)
Scincidae (skinks)						
Madascincus melanopleura	2	14	60	1	ND1	(73)

Table 3.2. Top-ranking models from our migrate-n analyses for each species. Visual descriptions of each model are provided in Fig. 3.3. For species with occurrences in the Montagne d'Ambre region, a secondary analysis was performed to assess whether Montagne d'Ambre represents a distinct population (Full M.d'A. Model) or a population panmictic with the northern highlands (M.d'A. Panmictic). The top-ranked model for each species are in bold. All log-marginal likelihood values were calculated using a Bézier curve approximation.

	Bayes Factor (Bézier-approximated log-marginal likelihood)					
	Full Model	**N/C Panmictic**	**C/S Panmictic**	**All Panmictic**	**Full M.d'A. Model**	**M.d'A. Panmictic**
Tenrecidae (tenrecs)						
Microgale drouhardi	**0 (-4662.33)**	-29.00 (-4676.83)	-114.70 (-4719.68)	-141.61 (-4733.13)	-	-
Microgale fotsifotsy	**0 (-3565.80)**	-88.95 (-3610.27)	-28.97 (-3580.28)	-125.13 (-3628.36)	**0 (-3688.14)**	-56.71 (-3716.50)
Microgale gymnorhyncha	**0 (-3397.12)**	-72.44 (-3433.33)	-23.75 (-3408.99)	-100.96 (-3447.59)	-	-
Microgale longicaudata	**0 (-2943.02)**	-134.00 (-3010.02)	-9.99 (-2948.02)	-147.62 (-3016.83)	**0 (-3547.57)**	-34.95 (-3565.04)
Microgale majori	**0 (-3797.91)**	-26.75 (-3811.29)	-24.78 (-3810.30)	-49.09 (-3822.45)	-	-
Microgale parvula	**0 (-2602.23)**	-18.24 (-2611.35)	-11.11 (-2607.78)	-32.61 (-2618.53)	**0 (-2643.36)**	-9.33 (-2648.03)
Microgale principula	**0 (-2130.88)**	-5.42 (-2133.59)	-17.97 (-2139.87)	-29.12 (-2145.44)	-	-
Microgale soricoides	**0 (-3096.25)**	-138.42 (-3165.46)	-14.22 (-3103.36)	-165.30 (-3178.90)	-	-
Microgale taiva	**0 (-2728.26)**	-	-	-34.47 (-2745.49)	-	-
Microgale thomasi	**0 (-1929.53)**	-0.02 (-1929.53)	-0.48 (-1929.77)	-1.30 (-1930.18)	-	-
Nesogale dobsoni	**0 (-2083.41)**	-7.21 (-2087.01)	-1.31 (-2084.06)	-11.74 (-2089.28)	-	-
Nesogale talazaci	-1.55 (-1947.11)	-9.98 (-1951.33)	**0 (-1946.34)**	-8.91 (-1950.80)	-0.86 (-1997.62)	**0 (-1997.19)**
Oryzorictes hova	**0 (-3197.16)**	-73.3 (-3233.81)	-26.5 (-3210.40)	-102.96 (-3248.63)	-	-
Nesomyidae (rodents)						
Eliurus grandidieri	**0 (-3226.20)**	-	-	-70.04 (-3261.22)	-	-
Eliurus majori	**0 (-4219.65)**	-69.25 (-4254.28)	-54.6 (-4246.96)	-130.56 (-4284.93)	**0 (-4381.15)**	-6.08 (-4384.19)
Eliurus minor	**0 (-5073.92)**	-53.7 (-5100.77)	-30.94 (-5089.39)	-93.8 (-5120.82)	-	-
Eliurus tanala	**0 (-3841.32)**	-262.31 (-3972.48)	-45.79 (-3864.22)	-350.56 (-4016.60)	**0 (-3890.85)**	-20.21 (-3900.96)
Eliurus webbi	**0 (-3401.57)**	-93.32 (-3448.23)	-41.13 (-3422.14)	-173.01 (-3488.08)	-	-
Gymnuromys roberti	**0 (-2995.04)**	-12.05 (-3001.06)	-4.78 (-2997.43)	-16.66 (-3003.37)	-	-
Nesomys rufus	**0 (-3339.74)**	-17.22 (-3348.34)	-3.91 (-3341.69)	-24.69 (-3352.08)	-	-
Gekkonidae (geckos)						
Hemidactylus mercatorius	**0 (-1118.20)**	-0.9 (-1118.65)	-4.02 (-1120.21)	-10.13 (-1123.26)	-1.17 (-1128.79)	**0 (-1128.21)**
Phelsuma lineata	**0 (-1083.41)**	-3.03 (-1084.92)	-3.47 (-1085.15)	-7.03 (-1086.92)	**0 (-1234.27)**	-2.3 (-1235.42)
Uroplatus sameiti	**0 (-2417.16)**	-2.07 (-2418.20)	-8.66 (-2421.49)	-24.56 (-2429.44)	-	-
Uroplatus sikorae	**0 (-2554.87)**	-17.82 (-2563.78)	-18.47 (-2564.11)	-43.55 (-2576.65)	**0 (-2761.86)**	-10.88 (-2767.30)
Chameleonidae (chameleons)						
Calumma nasutum	**0 (-2399.44)**	-34.86 (-2416.86)	-74.05 (-2436.46)	-113.74 (-2456.31)	-	-
Scincidae (skinks)						
Madascincus melanopleura	**0 (-2193.88)**	-114.09 (-2250.92)	-45.32 (-2216.54)	-161.78 (-2274.76)	**0 (-2253.92)**	-11.42 (-2259.62)

Table 3.3. Modal values from msBayes analyses, which were used to determine whether the divergence events we identified were temporally congruent across all taxa. HPD = highest posterior density interval; $E(\tau)$ = mean divergence time across taxon-pairs, in coalescent units; $var(\tau)$ = the variance in divergence time across taxon-pairs, in coalescent units; $\Omega = var(\tau) / E(\tau)$. If the HPD of Ω includes 0.0, we cannot reject simultaneous divergence.

Divergence Event	n(taxon pairs)	$E(\tau)$ [95% HPD]	$var(\tau)$ [95% HPD]	Ω [95% HPD]
Mt. d'Ambre – Northern Highlands	9	0.048 [0.000 - 0.153]	0.014 [0.000 - 0.034]	0.120 [0.031 - 0.257]
Northern Highlands – Central Highlands	24	0.204 [0.121 - 0.272]	0.008 [0.000 - 0.016]	0.036 [0.000 - 0.060]
Central Highlands – Southern Highlands	18	0.051 [0.000 - 0.106]	0.000 [0.000 - 0.008]	0.020 [0.000 - 0.047]

3.9 Figures

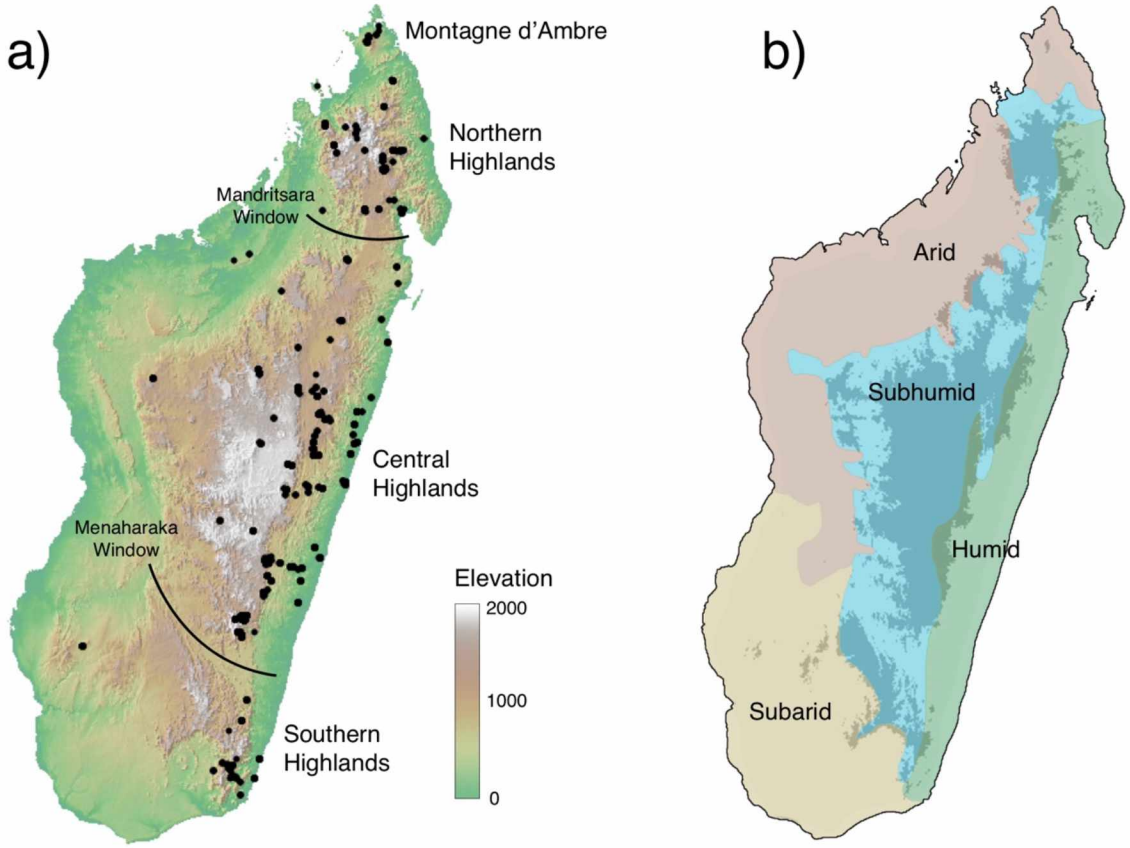


Figure 3.1. Maps of Madagascar showing (a) elevation and (b) simplified bioclimatic zones. All species evaluated in this study are endemic to and widespread throughout the eastern humid and subhumid forests. Black points on map (a) are collection localities, while curved lines denote the two low-elevation divisions discussed in the text. Darker areas in map (b) are >900 m elevation.

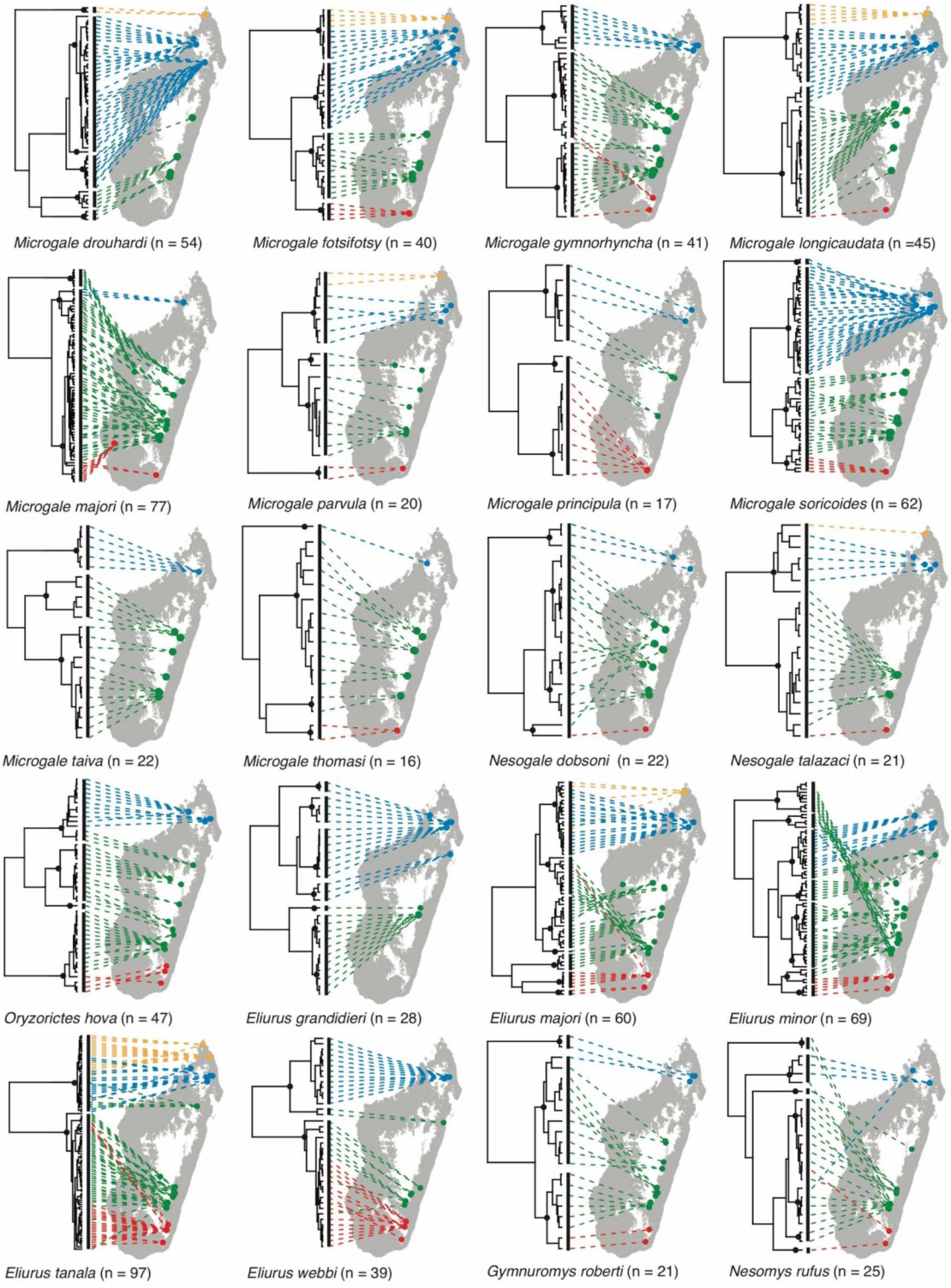


Figure 3.2. Phylogenetic trees and locality maps for each small-mammal species. Dashed lines connect each individual to its collection locality. Lines and points are colored according to high-elevation region: Montagne d'Ambre (yellow), northern highlands (blue), central highlands (green), and southern highlands (red). Vertical black lines to the right of the phylogeny denote the candidate species identified in our bGMYC analysis, while black points on nodes (or, in some cases, individual tips) denote the genetic clusters identified in our dapc analysis.

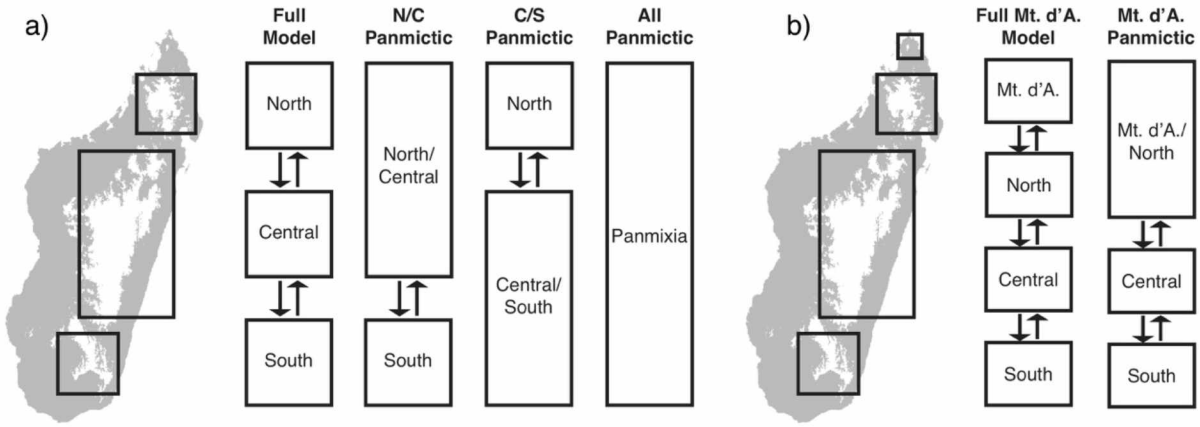


Figure 3.3. A schematic diagram of the models evaluated in our (a) first and (b) second rounds of Migrate-N analyses. In the full models, each high-elevation region (Montagne d'Ambre and the northern, central, and southern highlands) is treated as a separate population with migration allowed between adjacent regions. In models that include panmixia, individuals from adjacent regions are treated as a single population. Not all models were tested for all species, as some species were absent from one or more region. N = northern highlands, C = central highlands, S = southern highlands, and Mt. d'A. = Montagne d'Ambre.

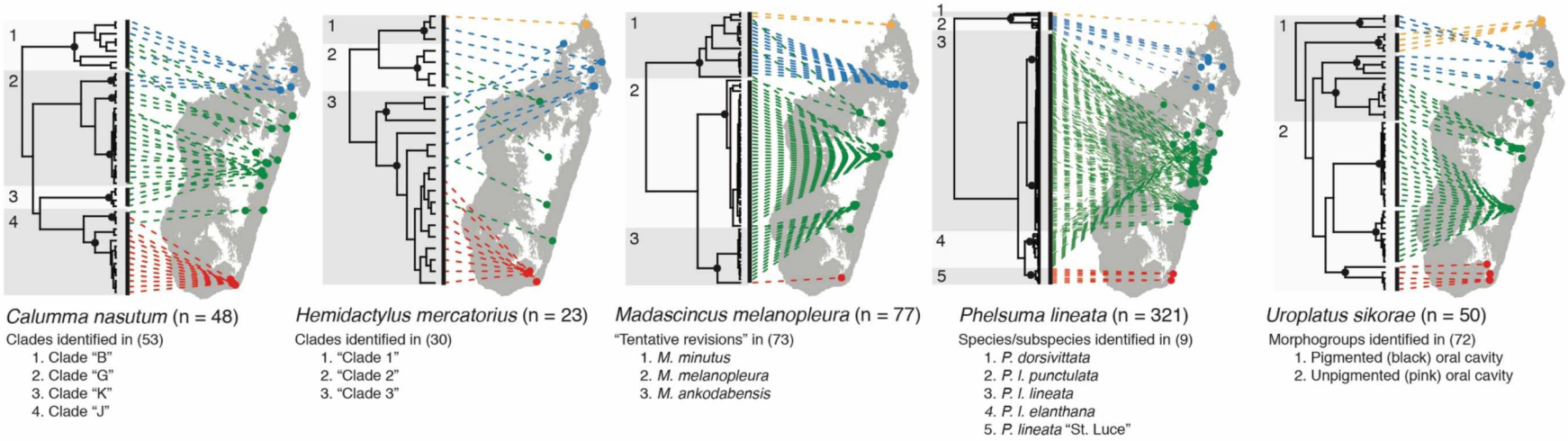


Figure 3.4. Phylogenetic trees and locality maps for each reptile species analyzed in this study. Numbered gray boxes overlaying each phylogeny show candidate species that have been identified in previous studies; these are associated with a written description below each species, with references shown in parentheses. All other colors and notations are as in Fig. 3.2.

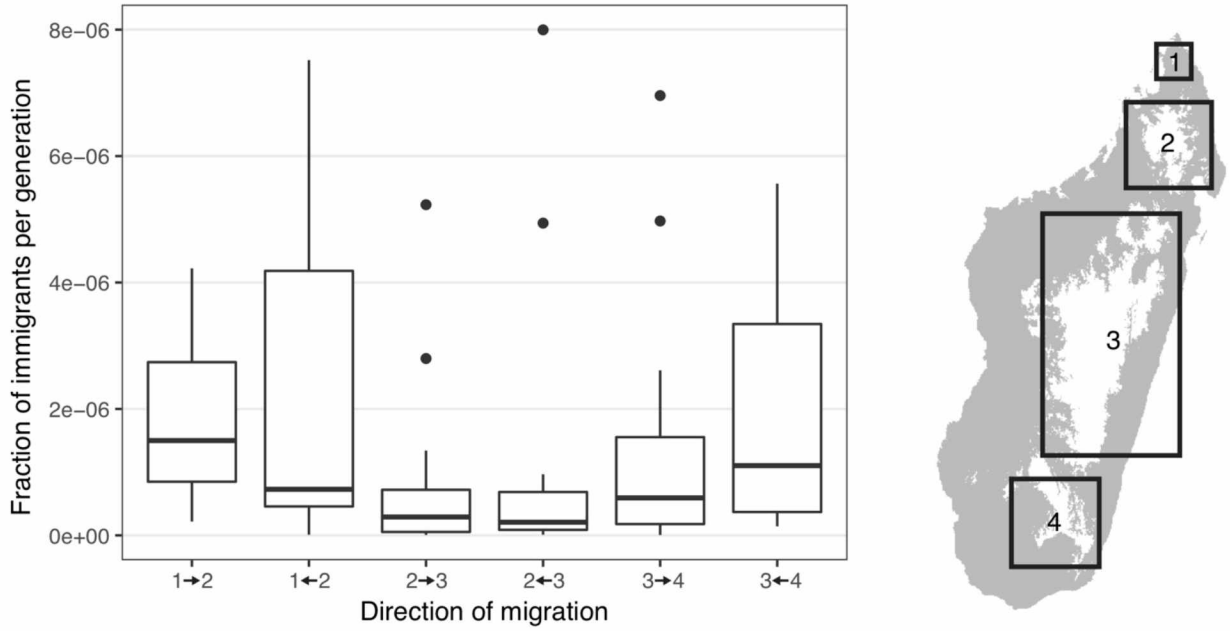


Figure 3.5. Boxplot of estimated migration rates between adjacent high-elevation regions. The direction of migration corresponds to the numbered boxes on the map (right).

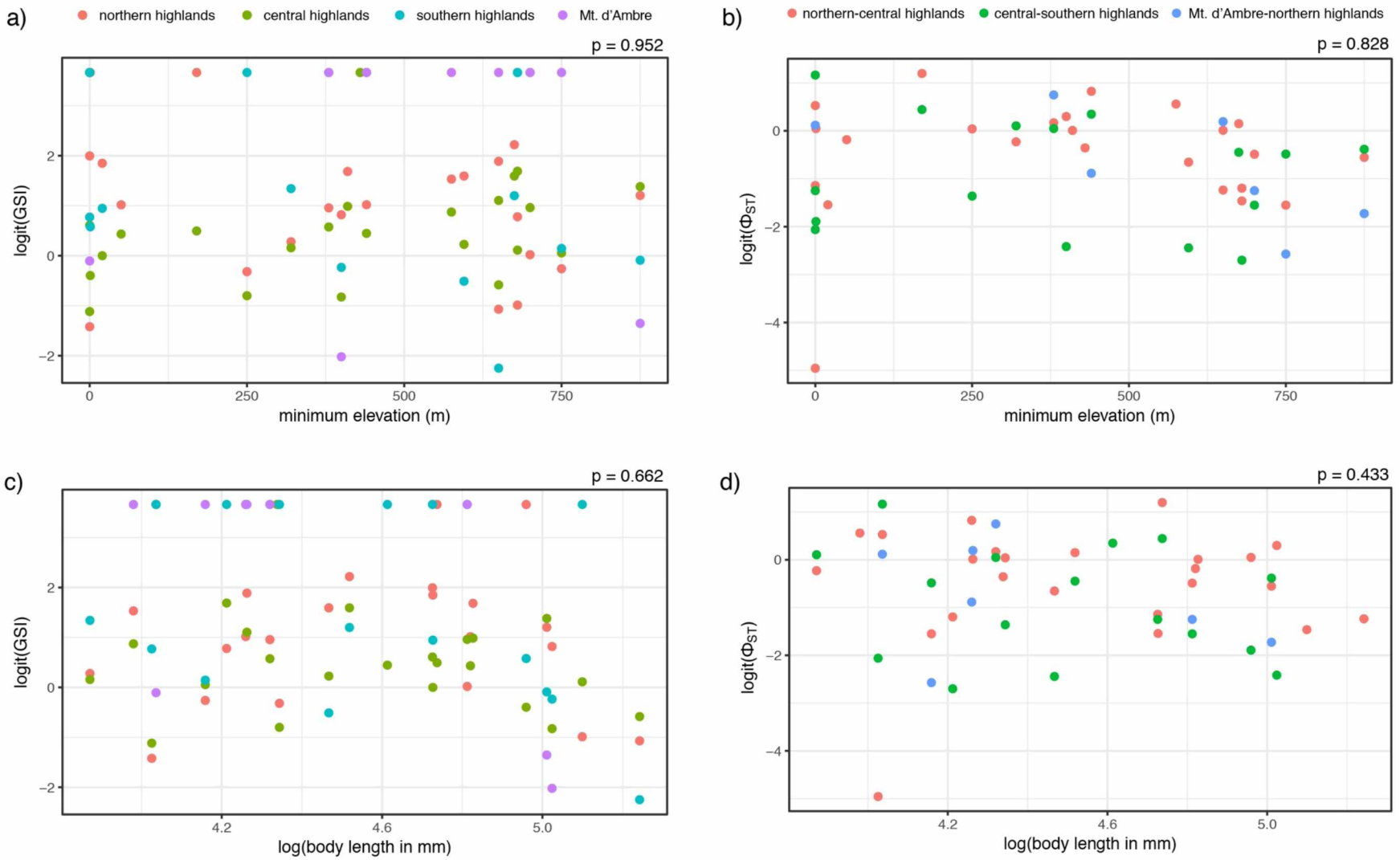


Figure 3.6. Linear regressions of the genealogical sorting index (GSI; a,c) and Φ_{ST} (b,d) for each species versus elevational range minimum (a,b) and body length (c,d). Body length does not include the tail, and is measured as head-body length in mammals and snout-vent length in reptiles.

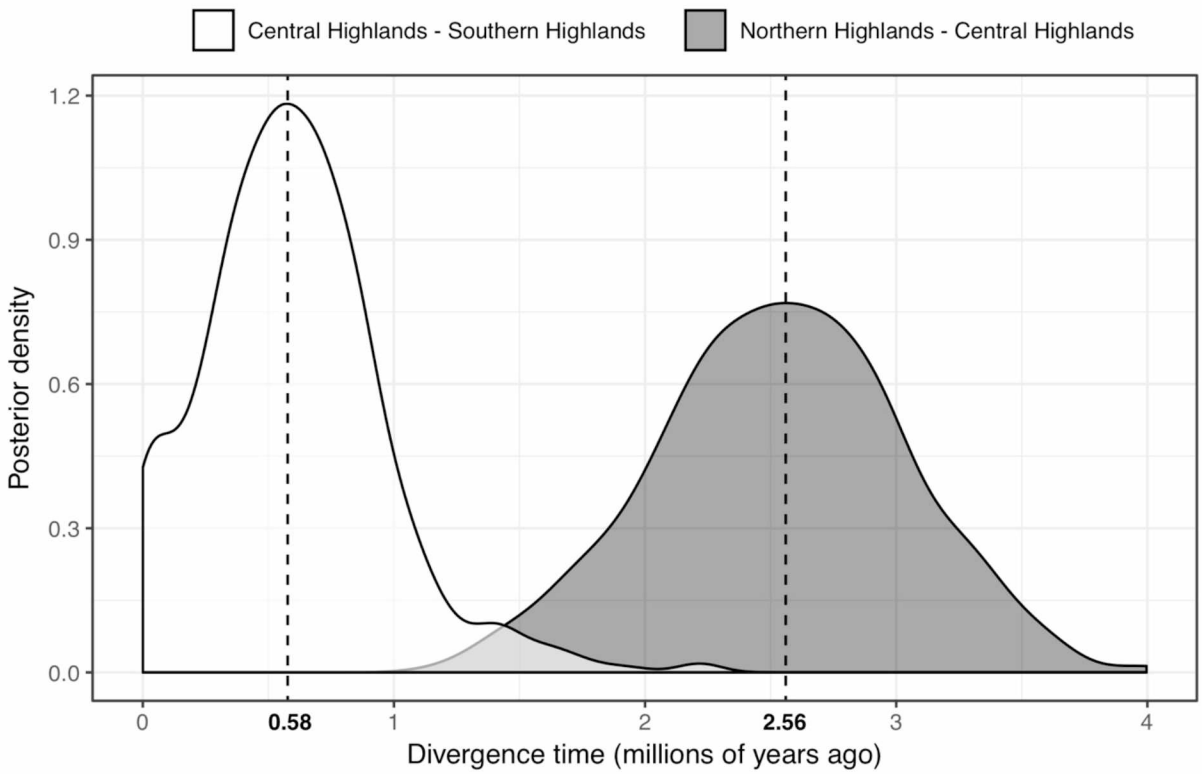


Figure 3.7. Posterior distributions of the mean estimated divergence times for the northern-central highlands divergence event and the central-southern highlands divergence event. Using an ABC approach (msBayes), we estimated synchronous divergence times across all evaluated taxa for both of these events.

3.10 Supplementary Tables

Supplementary Table 3.1. Catalog numbers, locality information, and bGMYC candidate species assignments for all small-mammal individuals included in this study. Collection abbreviations are as follows: Field Museum of Natural History (FMNH), Museum of Vertebrate Zoology at Berkeley (MVZ), Smithsonian Institution National Museum of Natural History (USNM), Université d'Antananarivo Département de Biologie Animale (UADBA), American Museum of Natural History (AMNH). This table is available upon request, and will be made publicly available upon this chapter's acceptance in a peer-reviewed journal.

Supplementary Table 3.2. GenBank accession numbers, locality information, and bGMYC candidate species assignments for all reptile species included in this study. These data were generated in previous studies. Some coordinates (shown in italics) were approximated using published locality descriptions. This table is available upon request, and will be made publicly available upon this chapter's acceptance in a peer-reviewed journal.

Supplementary Table 3.3. Table of all ecological data used in the multiple linear regression analyses. "min E" = minimum elevation, "GSI" = genealogical sorting index. "Body length" is given as head-body length (for mammals) or snout-vent length (for reptiles).

Species	min E (m)	GSI (Mt. d'Ambre)	GSI (northern highlands)	GSI (central highlands)	GSI (southern highlands)	Average GSI	Taxonomic group	Body length (mm)
Microgale drouhardi	650	1.0000	0.8879	0.7644	N/A	0.8841	tenrec	71
M. fotsifotsy	380	1.0000	0.7341	0.6471	1.0000	0.8453	tenrec	75.2
M. gymnorhyncha	595	N/A	0.8485	0.5592	0.3684	0.5920	tenrec	87.1
M. longicaudata	440	1.0000	0.7471	1.0000	N/A	0.9157	tenrec	70.8
M. majori	680	N/A	0.6952	0.8623	1.0000	0.8525	tenrec	67.5
M. parvula	0	0.4722	1.0000	1.0000	1.0000	0.8681	tenrec	56.6
M. principula	250	N/A	0.4167	0.3000	1.0000	0.5722	tenrec	77
M. soricoides	675	N/A	0.9231	0.8485	0.7826	0.8514	tenrec	91.7
M. taiva	430	N/A	1.0000	1.0000	N/A	1.0000	tenrec	76.6
M. thomasi	440	N/A	N/A	0.6154	1.0000	0.8077	tenrec	100.8
Nesogale dobsoni	170	N/A	1.0000	0.6275	N/A	0.8137	tenrec	114.1
Nesogale talazaci	50	N/A	0.7467	0.6122	N/A	0.6795	tenrec	124
Oryzorictes hova	20	N/A	0.8831	0.5000	0.7321	0.7051	tenrec	112.9
Eliurus grandidieri	410	N/A	0.8615	0.7404	N/A	0.8010	nesomyine	124.8
E. majori	875	0.1896	0.7834	0.8150	0.4763	0.5661	nesomyine	150
E. minor	0	N/A	0.9003	0.6553	1.0000	0.8519	nesomyine	112.8
E. tanala	400	0.0968	0.7042	0.2941	0.4386	0.3834	nesomyine	152
E. webbi	1	N/A	1.0000	0.3968	0.6481	0.6817	nesomyine	142.5
Nesomys rufus	650	N/A	0.2424	0.3506	0.0741	0.2224	nesomyine	189.1
Gymnuromys roberti	680	N/A	0.2593	0.5294	1.0000	0.5962	nesomyine	163.9
Calumma nasutum	320	N/A	0.5727273	0.5414634	0.8081633	0.6408	reptile	48
Hemidactylus mercatorius	0	N/A	0.1786	0.2333	0.6933	0.3684	reptile	56
Madascincus melanopleura	575	1.0000	0.8392	0.7162	N/A	0.8518	reptile	53.5
Phelsuma lineata	750	1.0000	0.4317	0.5145	0.5380	0.6211	reptile	64
Uroplatus sikorae	700	1.0000	0.5051	0.7351	1.0000	0.8100	reptile	123

Supplementary Table 3.4. Information used to calculate average mutation rates and mutation scalars (relative mutation rates) which were then used in the msBayes simultaneous divergence analyses. The average number of pairwise mutations was calculated from sequence data.

Species 1	Species 2	Locus	Divergence Time (MYA)	Avg. # pairwise mutations	Mutation rate	Relative Rate	Divergence time reference
Calumma nasutum	*Calumma boettgeri*	ND2	4.32	44.556	1.02E-08	5.24	Zheng & Wiens 2016
Eliurus grandidieri	*Brachytarsomys albicauda*	CytB	15	224.786	6.24E-09	3.21	Poux et al. 2005
Hemidactylus mercatorius	*Hemidactylus mabouia*	16S	2.92	13.138	4.45E-09	2.29	Zheng & Wiens 2016
Microgale fotsifotsy	*Microgale soricoides*	ND2	6	130.063	1.04E-08	5.34	Everson et al. 2016
Phelsuma lineata	*Phelsuma kely*	16S	8.57	20.619	2.41E-09	1.24	Zheng & Wiens 2016
Uroplatus sameiti	*Uroplatus sikorae*	ND4	8	40.835	3.65E-09	1.88	Raxworthy et al. 2008
Madascincus melanopleura	*Amphiglossus macrocercus*	ND1	44.17	81.026	1.94E-09	1.00	Zheng & Wiens 2016

References

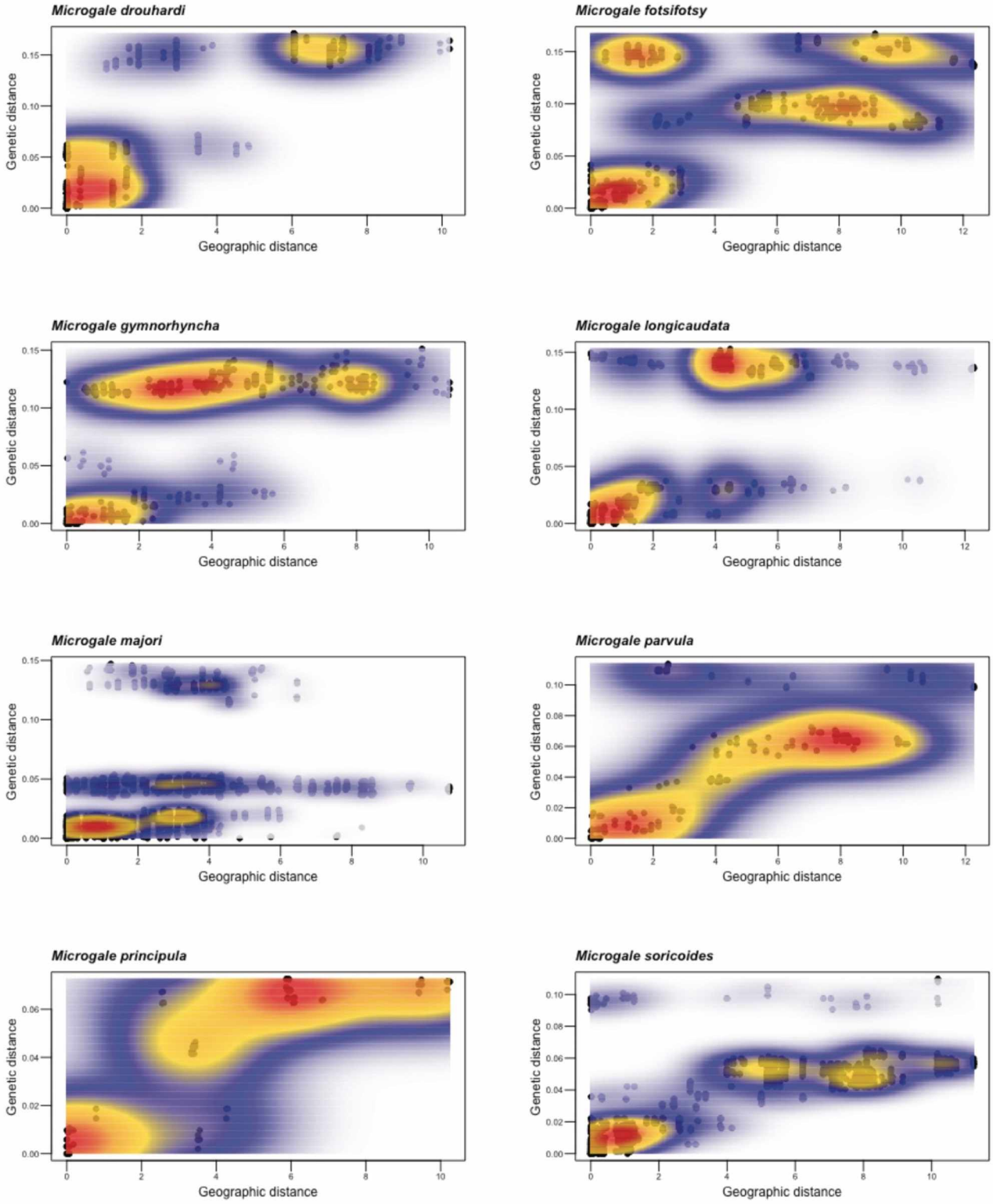
Zheng, Y., & Wiens, J. J. (2016). Combining phylogenomic and supermatrix approaches, and a time-calibrated phylogeny for squamate reptiles (lizards and snakes) based on 52 genes and 4162 species. Molecular phylogenetics and evolution, 94, 537-547.

Poux, C., Madsen, O., Marquard, E., Vieites, D. R., de Jong, W. W., & Vences, M. (2005). Asynchronous colonization of Madagascar by the four endemic clades of primates, tenrecs, carnivores, and rodents as inferred from nuclear genes. Systematic Biology, 54, 719-730.

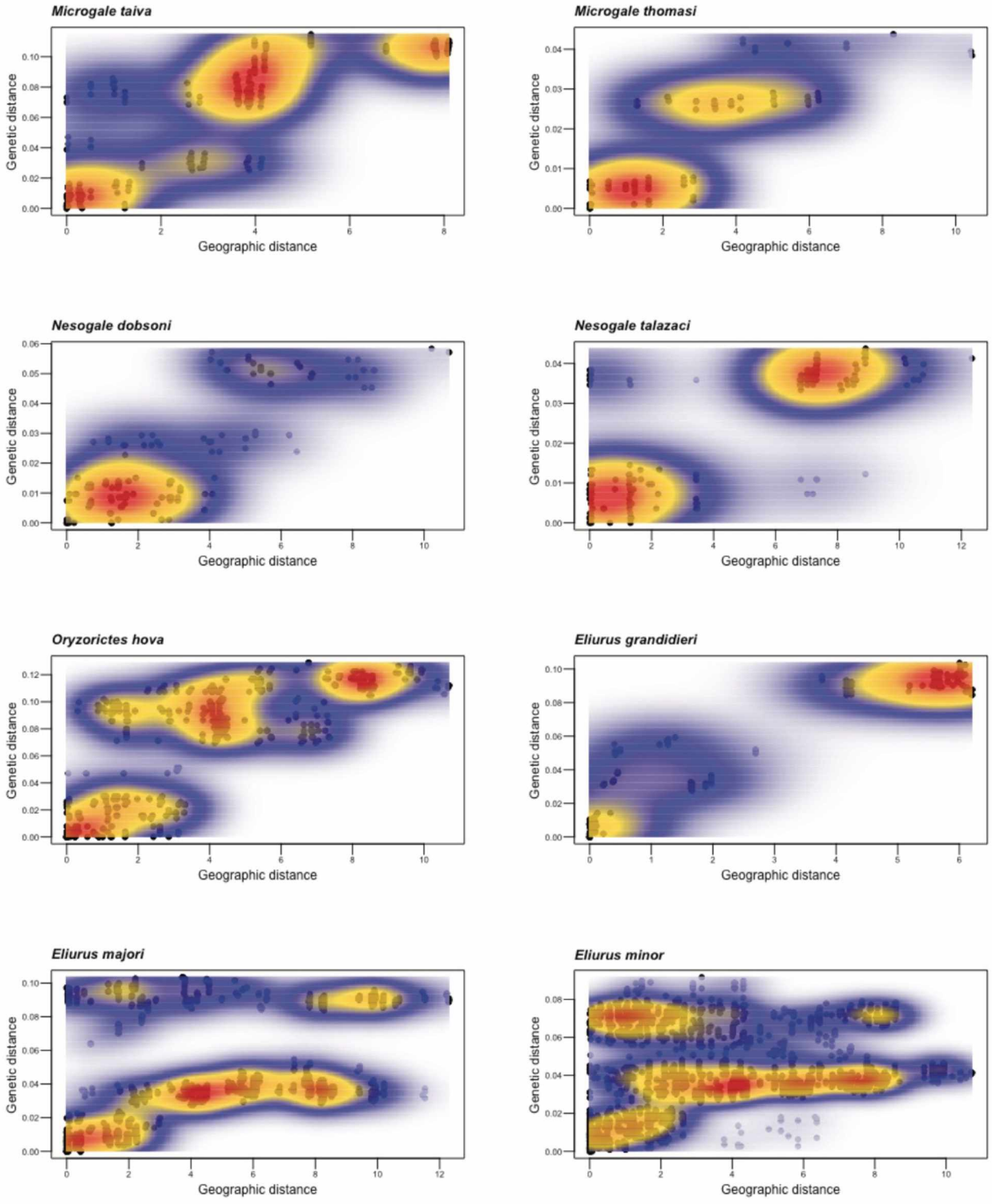
Everson, K. M., Soarimalala, V., Goodman, S. M., & Olson, L. E. (2016). Multiple loci and complete taxonomic sampling resolve the phylogeny and biogeographic history of tenrecs (Mammalia: Tenrecidae) and reveal higher speciation rates in Madagascar's humid forests. Systematic biology, 65, 890-909.

Raxworthy, C. J., Pearson, R. G., Zimkus, B. M., Reddy, S., Deo, A. J., Nussbaum, R. A., & Ingram, C. M. (2008). Continental speciation in the tropics: contrasting biogeographic patterns of divergence in the Uroplatus leaf-tailed gecko radiation of Madagascar. Journal of Zoology, 275, 423-440.

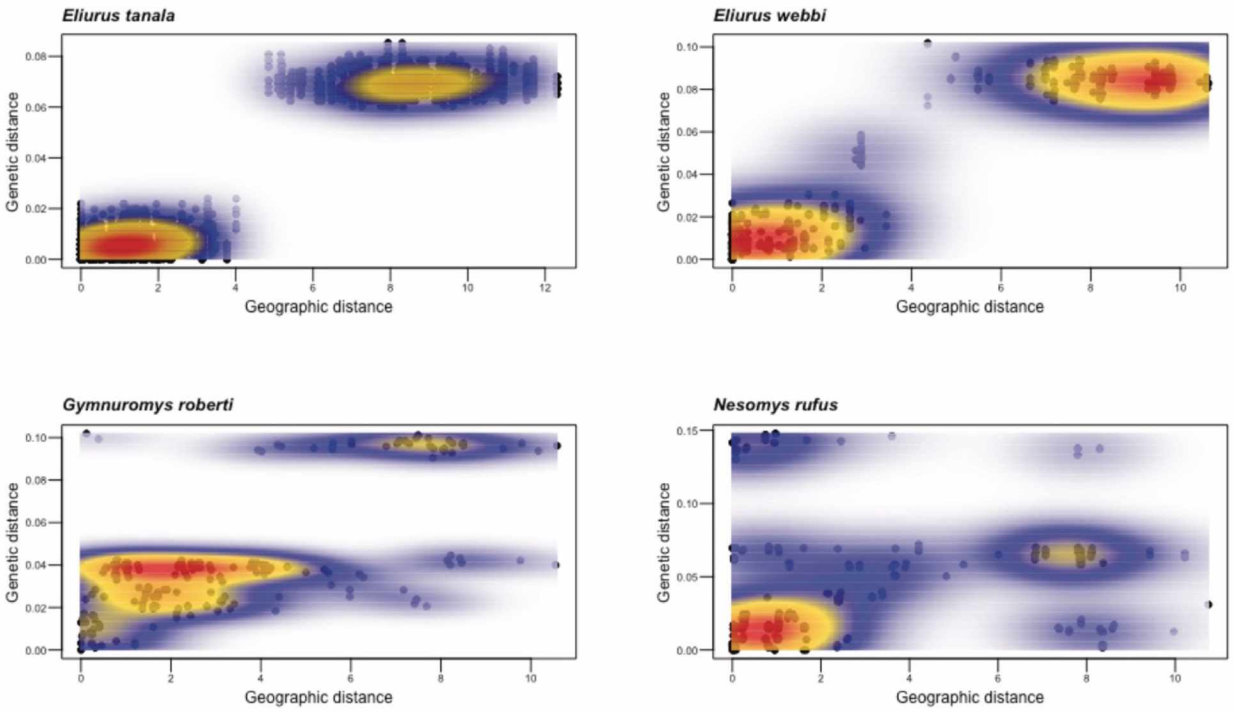
3.11 Supplementary Figures



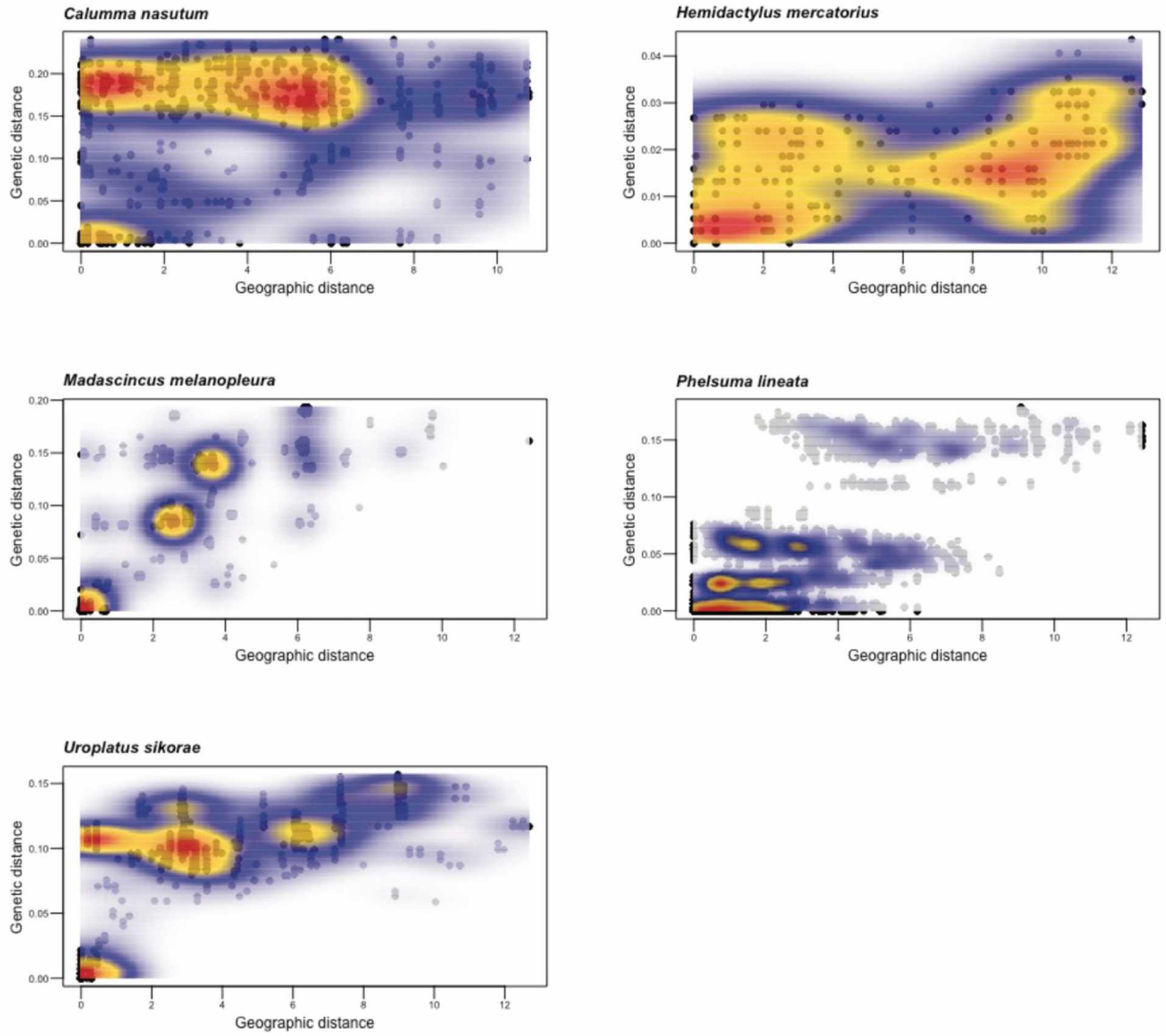
Supplementary Figure 3.1. Plots of genetic distance by geographic distance for all small mammals evaluated in this study, with point density heatmaps to aid in visualization of genetic clusters.



Supplementary Figure 3.1 (continued).



Supplementary Figure 3.1 (continued).



Supplementary Figure 3.2. Plots of genetic distance by geographic distance for all reptiles evaluated in this study, with point density heatmaps to aid in visualization of genetic clusters.

Chapter 4 – Speciation and gene flow in two sympatric shrew tenrecs, *Microgale fotsifotsy* and *M. soricoides* (Mammalia: Tenrecidae)[4]

[4] Everson, K.M., Olson, L.E. In preparation for *Molecular Ecology*.

4.1 Abstract

Madagascar's shrew tenrecs (Mammalia: Tenrecidae: *Microgale*) represent an excellent system for studying speciation. Most are endemic to, and broadly sympatric throughout, the island's humid-forest belt, a region renowned for high levels of biodiversity and high rates of *in situ* diversification. Here we use a combination of skeletal measurements, mitochondrial DNA sequence data, and next-generation nuclear sequence data from two species of shrew tenrecs, *M. fotsifotsy* and *M. soricoides*, to better understand their speciation history. Our mitochondrial DNA-derived phylogeny recovered two distinct clades of *M. fotsifotsy* that are not reciprocally monophyletic, with one occuring only in the far north of Madagascar and the other far more widespread and broadly sympatric with *M. soricoides*. Nuclear data corroborate these two clades and demonstrate that gene flow has occurred only between *M. soricoides* and the widespread clade of *M. fotsifotsy*. Morphometric data reveal major differences between *M. soricoides* and *M. fotsifotsy*, as well as more subtle but significant differences between the two clades of *M. fotsifotsy*. Our findings support the idea that isolation and re-expansion from refugia played a role in the diversification of Madagascar's humid-forest taxa, and that species diversity in shrew tenrecs is greater than currently recognized.

4.2 Introduction

Speciation is often described as a continuum along which many changes gradually accumulate as two lineages diverge on the path to reproductive isolation. Studying taxa at the early stages of this process poses unique challenges; for example, morphological differences might be subtle, absent, or not yet fixed (Bickford *et al.* 2007; Sáez & Lozano 2005). In such cases, biodiversity is likely to be underestimated as so-called cryptic species go unrecognized (e.g., Vieites *et al.* 2009). Furthermore, a growing body of evidence has shown that speciation can occur in spite of ongoing gene flow (Nosil 2008; Pinho & Hey 2010), either during the speciation process (speciation with gene flow) or after a period of isolation (secondary contact). Because most analytical programs that estimate phylogenetic and demographic histories do not account for gene flow, this can produce spurious estimates of population sizes, divergence times, and phylogenetic relationships (Eckert & Carstens 2008; Leaché *et al.* 2014; Morales *et al.* 2016).

Madagascar is a natural laboratory for speciation research. This island off the southeast coast of Africa has been isolated for >88 million years, during which time numerous taxonomic groups have colonized the island and undergone *in situ* radiation (Vences *et al.* 2009). One such native lineage is the family Tenrecidae, or the tenrecs, which are small- to medium-sized placental mammals that first colonized Madagascar ~30-56 MYA (Everson *et al.* 2016) and have been touted as a "remarkable adaptive radiation" (Eisenberg & Gould 1970, p. 1). Tenrec taxonomy is rapidly changing, with eight extant and one extinct species described and another resurrected from synonymy in the last 25 years (e.g.,

Goodman et al. 2006; Goodman & Soarimalala 2004; Olson et al. 2009), plus recent studies suggesting the existence of additional cryptic species (Everson *et al.* 2018; Olson *et al.* 2004; Poux *et al.* 2008). Recent species discoveries are largely due to three factors: scientific collecting in previously undocumented areas of the island; increased use of pitfall traps, which have proven to be particularly effective for tenrecs (S.M. Goodman, *pers. comm.*); and the application of new molecular and morphometric techniques for discovering and delimiting species (e.g., Blaimer & Fisher 2013; Miralles & Vences 2013; Olivieria *et al.* 2007; Vieites *et al.* 2009). Documenting previously unrecognized species continues to be urgently important in light of Madagascar's exceptionally high rates of habitat loss (Grinand *et al.* 2013; Harper *et al.* 2007).

Shrew tenrecs (*Microgale* spp.) are the most speciose genus in the family Tenrecidae, encompassing 21 of the 32 currently recognized tenrec species. As their common name implies, shrew tenrecs are small-bodied (2-49 g) and shrew-like in appearance, with some species having adaptations for digging, swimming, or climbing (Jenkins 2003). Most shrew tenrecs occur in the humid forests that form a north-south belt along the eastern side of the island, where fourteen of these species have remarkably similar, widespread distributions (Fig. 4.1). Given their apparently overlapping niche space, it is not well understood why the shrew tenrecs became so speciose relative to other tenrec clades. It has been proposed that speciation in Madagascar's humid-forest taxa has been primarily allopatric, driven by climatic conditions that have caused forest habitats to be fragmented and/or restricted to high elevations (Everson *et al.* 2018; Wilmé *et al.* 2006; Wollenberg *et al.* 2008). Recent advances in the collection and analysis of genetic data allow certain aspects of the speciation process to be reconstructed in great detail, including the timing of divergence and the presence or absence of gene flow among lineages (Gutenkunst *et al.* 2009; Hey 2010; Jackson *et al.* 2016). If speciation was allopatric, we would expect to find genetic signatures of isolation, possibly with secondary contact after range re-expansion; conversely, sympatric speciation would likely produce a genetic signature of gene flow during the earliest stages of speciation (Coyne & Orr 2004; Pinho & Hey 2010).

Here we focus on two species of shrew tenrecs, *M. fotsifotsy* Jenkins *et al.* 1997 and *M. soricoides* Jenkins 1993, which have nearly identical, widespread distributions throughout Madagascar's eastern humid forests and are often collected from the same localities (Fig. 4.2). Recent molecular evidence (Everson *et al.* 2016) has shown that *M. fotsifotsy* and *M. soricoides* are closely related, forming a clade with one other species, *M. nasoloi* Jenkins & Goodman 1999, which is known from just four localities in western dry forest habitats. All three species are poorly studied but thought to be insectivorous and terrestrial-scansorial. The goal of this study is to use integrative data to better understand the speciation process in this clade, including the presence or absence of speciation-with-gene-flow, secondary contact, and morphological divergence. We use a combination of mitochondrial DNA

sequence data, nuclear DNA sequence data, and skeletal measurements to estimate phylogenetic relationships and gene flow between species, and to re-assess current taxonomic designations.

4.3 Materials and Methods

DNA sequence data collection

Fresh or frozen tissue samples were obtained from three different museum collections for 40 vouchered specimens of *Microgale fotsifotsy*, 62 specimens of *M. soricoides*, one specimen of *M. nasoloi*, and two specimens of *M. cowani* (outgroup) from 20 collection localities (Supplementary Table 4.1). We followed DNA extraction, amplification, and sequencing protocols outlined in Everson et al. (2016) to collect mitochondrial sequence data (mtDNA) from the NADH dehydrogenase subunit 2 (ND2) gene.

A subset of the DNA extractions used in our mtDNA analyses (9 specimens of *M. fotsifotsy*, 7 of *M. soricoides*, and 2 of *M. cowani*) were also used to collect nuclear DNA (nuDNA) sequence data (Supplementary Table 4.2) via the ultraconserved elements (UCE) sequence-capture protocol (Faircloth *et al.* 2012). We quantified the amount of double-stranded DNA in each extraction using a Qubit 2.0 fluorometer (Life Technologies) and visualized the molecular weight on an agarose gel using a genomic-weight ladder. We selected the subset of extractions to include individuals with large amounts of high-molecular-weight DNA that also represented several different localities across each species' range. The sole tissue extraction from *M. nasoloi* did not contain a sufficient amount of high-molecular-weight DNA for this sequence-capture protocol.

We fragmented the DNA to approximately 600 base pairs using a Bioruptor sonication instrument (Diagenode). TruSeq genomic libraries were prepared using KAPA library kits (Kapa Biosystems) and iTru adapters (Faircloth & Glenn 2012), then pooled in pre-enrichment libraries of eight samples each. Ultraconserved elements were enriched using a set of 5,472 probes previously designed for use in tetrapods (Faircloth *et al.* 2012). Probes were manufactured by MYcroarray, Inc. (Ann Arbor, USA), and we followed their enrichment protocol. Post-enrichment libraries were pooled at equimolar ratios and were sequenced on a portion of a lane (alongside other samples unrelated to this project) of an Illumina HiSeq2000 instrument with a 150-bp paired-end flow cell at the University of California Los Angeles Neuroscience Genomics Core.

We used the PHYLUCE bioinformatics pipeline (Faircloth 2016), including a recently developed SNP phasing workflow (Andermann *et al.* 2018), to post-process the sequencing reads and generate aligned data. Reads were first trimmed for adapter contamination and low-quality bases using Illumiprocessor (Bolger *et al.* 2014; Faircloth 2013) then assembled using Trinity v2.0 (Grabherr *et al.* 2011). We isolated UCE loci from the resulting assemblies and discarded putative paralogs using PHYLUCE python scripts. Alleles were then mapped onto UCE contigs using BWA-MEM (Li 2013),

sorted and phased using SAMTOOLS (Li 2011), and cleaned using PICARD (Broad Institute, available from http://broadinstitute.github.io/picard). Cleaned, phased contigs were aligned and edge-trimmed using MAFFT (Katoh & Standley 2013). We created a 75% complete dataset (wherein each UCE locus contains data from at least 75% of individuals) and produced locus-specific and concatenated Nexus alignment files using PHYLUCE scripts. We also generated SNP-only alignment files by extracting biallelic sites from sequence alignments using the python function snps_from_uce_alignments.py (Andermann et al. 2018; available from github.com/tobiashoffmann88).

Phylogenetic analyses and divergence time estimation

We used MrBayes v3.2.0 (Huelsenbeck & Ronquist 2001) to estimate two phylogenies via Bayesian Inference (BI): a phylogeny of the mitochondrial ND2 gene containing 105 individuals, and a phylogeny of 2,836 concatenated UCE loci containing 18 individuals. Both analyses were conducted using two independent MCMC runs (four chains each) of 10,000,000 iterations each, sampling trees every 1000 iterations. We combined the runs using LogCombiner (Drummond & Rambaut 2007) and assessed stationarity by evaluating the likelihood scores of the MCMC chains in Tracer v1.6 (Rambaut & Drummond 2007). The first 20% of all trees were removed as burn-in, and the last 8,000 trees were used to construct a majority-rule consensus tree and assign Bayesian posterior probabilities (BPPs) to each node.

We also estimated a species phylogeny using a summary-based coalescent approach. Summary-based methods require well-resolved gene trees as input (Xi *et al.* 2015), which may not be produced by UCE data that contain reduced phylogenetic signal. To address this potential problem, we used a PHYLUCE script (phyluce_align_get_informative_sites) to count the number of informative sites in each locus of our UCE dataset, which we then used to isolate only the loci with informative sites in the upper quartile of the range. We estimated maximum-likelihood gene trees for each of these most-informative UCE loci using RAxML (Stamatakis 2014) under the GTR-GAMMA site-rate substitution model. Resulting gene trees were used as input to the summary-based species-tree program ASTRAL v4.10.12 (Mirarab *et al.* 2014).

We then used SNP data from our UCE loci to estimate a species tree via the multispecies coalescent model in SNAPP (Bryant *et al.* 2012), implemented in BEAST2 (Bouckaert *et al.* 2014). SNP data were thinned to include only biallelic sites, then converted to an XML file using BEAUti, part of the BEAST2 package. We then modified this XML file using the snapp_prep.rb script from Stange et al. (2018) to enable divergence time estimation. We estimated divergence times by placing normally distributed priors at two nodes, based on estimated dates from Everson et al. (2016): the divergence of *M. fotsifotsy* and *M. soricoides* (mean 6.0 MYA, standard deviation 1.0), and the divergence of *M. cowani*

from the remaining taxa in this study (mean 12.6 MYA, standard deviation 1.0). We ran BEAST2 for 1 million generations, sampling every 2,000 generations. Tracer was used to view the MCMC output and check for convergence, and TreeAnnotator was used to calculate node support values and node ages after discarding a 20% burn-in.

We also estimated the change in population sizes over time using extended Bayesian skyline plots (EBSPs). As the analysis of all UCE loci would be computationally prohibitive, we performed this analysis using only the 50 UCE loci with the greatest numbers of informative sites. We performed separate analyses for *M. soricoides* and each of the two clades of *M. fotsifotsy* (discussed below) using BEAST 1.8.3 (Drummond & Rambaut 2007) with the same nucleotide substitution models and MCMC parameters used in the previous BEAST analysis. We checked for stationarity using Tracer and plotted EBSPs in R (R Core Team 2012) using the package ggplot2 (Wickham 2016).

Analyses of admixture and gene flow

We used our SNP dataset in the BI program fastStructure (Raj *et al.* 2014) to determine the number of genetic clusters in our dataset and individual admixture proportions. We performed analyses using a number of genetic clusters (K) from 1 to 10, then selected the best value of K using maximum marginal likelihoods calculated using the python script chooseK.py. Genetic clusters were visualized using DISTRUCT (Rosenberg 2003).

Our mitochondrial and nuclear phylogenetic analyses recovered different topologies with respect to *M. soricoides* and the two primary clades of *M. fotsifotsy* (see below). To test the idea that gene flow was responsible for these discordant topologies, we used an approximated likelihood framework implemented in the program PHRAPL (Jackson *et al.* 2017) to evaluate different models of gene flow between *M. fotsifotsy* and *M. soricoides*. Briefly, we used PHRAPL to simulate gene-tree distributions under 16 different phylogeographic models representing all possible bi-directional gene flow scenarios, then used the Akaike Information Criterion (AIC) to select the model that best fit our empirical gene trees (estimated using RaxML, described above). The 16 phylogeographic models varied in their inclusion or exclusion of pairwise symmetrical migration between groups on the species tree topology estimated in SNAPP (Supplementary Fig. 4.1). We simulated 100,000 trees for each model using the default mutation-scaled parameter values for divergence time (τ = 0.3, 0.58, 1.11, 2.12, 4.07, 7.81, and 15) and migration (m = 0.1, 0.22, 0.46, 1, 2.15, and 4.64).

Morphometric data collection and analyses

One of us (L.E.O.) recorded 23 cranial measurements, 9 mandibular measurements, and (where complete skeletons were available) 13 postcranial measurements from 126 vouchered adult specimens: 46

specimens of *Microgale fotsifotsy*, 79 of *M. soricoides*, and 1 of *M. nasoloi* (see Table 4.1 for measurement descriptions and Supplementary Table 4.3 for morphometric data). Many of these specimens were also included in our sequencing dataset (see Supplementary Tables 4.1 and 4.2). Adults were identified by the presence of a fully erupted permanent dentition and were measured with digital calipers accurate to 0.1 mm, with measurements of bilateral elements taken from the left side whenever possible. These measurements include 19 described in previous tenrec studies (Everson *et al.* 2018; S.M. Goodman *et al.* 2006; Olson *et al.* 2004, 2009) and four new to this study (OP, P4M3, Ci1, and PelvisGrLength; Table 4.1). These measurements were added to three external body measurements (head-body length, tail length, and hindfoot length) taken at the time of specimen collection, as in previous studies (e.g., Goodman & Jenkins 2000). We also calculated seven indices (measurement ratios) used in previous studies to infer locomotory and substrate-use adaptations (Salton & Sargis 2008, 2009; Woodman & Stabile 2015).

We confirmed that measurements within each species conformed to a normal distribution using a Shapiro-Wilk test, and confirmed that there were no differences between sexes in each species using a one-way analysis of variance (ANOVA) on each measurement or index; both tests were conducted in R (R Core Team 2012). We then used an ANOVA on each measurement or index to test for differences among species and clades. We also used R to perform a principal components analysis (PCA) to determine the major sources of variation across *M. fotsifotsy*, *M. soricoides*, and *M. nasoloi,* and performed a canonical variates analysis (CVA) to determine the most useful traits for predicting species assignments. We recognized four data modules (cranial, mandibular, postcranial, and external measurements) which were analyzed both separately and together. For the PCA and CVA, missing data were imputed using Multivariate Imputation by Chained Equations in the R package mice (van Buuren & Groothuis-Oudshoorn 2010). Specimens without vouchered postcranial skeletons were removed prior to imputation from all analyses that included the postcranial dataset.

4.4 Results

Sequence data collection

We generated 1,044 base pairs (bp) of mitochondrial (ND2) sequence data for 105 individuals and nuclear sequence data for 18 individuals using sequence capture of UCEs. After filtering and alignment, the UCE dataset included 2,836 phased loci (2,006,999 bp total) and 799 biallelic SNPs (Supplementary Table 4.2).

Phylogenetic analyses and divergence time estimation

Using mtDNA, we recovered a well-resolved phylogeny with two distinct clades of *M. fotsifotsy*: one that is restricted to the far north of Madagascar in the vicinity of Montagne d'Ambre ("north clade" hereafter) and one that is broadly distributed throughout the northern, central, and southern highlands and sympatric with *M. soricoides* ("widespread clade", Fig. 4.2). The north clade was recovered as sister to the geographically distant *M. nasoloi*, whereas the widespread clade was recovered as sister to *M. soricoides*.

Our analysis of concatenated UCE sequence data (which did not include *M. nasoloi*; see above) recovered the same two clades of *M. fotsifotsy* (Fig. 4.3), and our two species-tree analyses (ASTRAL and SNAPP) recovered the same basic topology as the concatenated analysis (Fig. 4.4). The divergence of the two *M. fotsifotsy* clades is estimated to have occurred 3.14 MYA [95% CI: 2.54-3.62 MYA] while the divergence of *M. soricoides* occurred 5.2 MYA [95% CI: 4.52 – 6.02 MYA]. We were not able to include nuclear data from *M. nasoloi* in these analyses, so we cannot confirm its phylogenetic placement; however, our mtDNA results, combined with a comparison of the divergence times estimated in Everson et al. (2016) to the divergence times estimated in our SNAPP analysis, suggest that this species is sister to the north clade of *M. fotsifotsy*.

Our extended Bayesian skyline plots generally showed a recent increase in population size across all species (Supplementary Fig. 4.2); however, all plots show large confidence intervals during the most recent time interval.

Analyses of admixture and gene flow

FastStructure identified three genetic clusters in our dataset. These groups corresponded to *M. soricoides*, *M. fotsifotsy* (north clade) and *M. fotsifotsy* (widespread clade). All individuals had <1% estimated admixture.

We used PHRAPL to simulate genetic data under 16 demographic models, and used AIC to select the model (or models) with the best fit to our empirical dataset. Model 14, which included gene flow between the two *M. fotsifotsy* clades, between *M. soricoides* and the widespread *M. fotsifotsy* clade, and between ancestral *M. fotsifotsy* and *M. soricoides*, was the top-ranking model (Fig. 4.5). The next-highest model (model 13) was nearly identical, but did not include gene flow between the two *M. fotsifotsy* clades. The estimated parameter values for model 14 fell at the lower ends of range of values we simulated: $\tau_{M.\,fotsifotsy\,-\,M.\,soricoides} = 7.81$, $\tau_{M.\,fotsifotsy\,\text{north}\,-\,\text{widespread}} = 1.11$, and $m = 0.1$. We do not know the substitution rates of ultraconserved elements in these species, nor do we know their generation times; however, if we assume a mutation-scaled theta of 0.004 (the median estimated value across all groups in our EBSP analyses), these parameter values are approximately equivalent to: divergence time of *M.*

'fotsifotsy and *M. soricoides* = 1.84 MYA, divergence of the two *M.'fotsifotsy* clades = 0.26 MYA, and proportion of migrants in each population per generation = 0.0004.

Morphometric data analyses

Principal components analyses showed a clear separation of *M. soricoides, M. nasoloi,* and *M.'fotsifotsy* in morphospace, although the two clades of *M.'fotsifotsy* were overlapping (Fig. 4.6, Supplementary Fig. 4.3). Similarly, our CVAs clearly differentiated *M. soricoides* from *M.'fotsifotsy*, but could not accurately discriminate the two *M.'fotsifotsy* clades from one another (Table 4.2). *Microgale nasoloi* could not be included in the CVAs as we only had measurement data from a single individual. All of the measurements we evaluated were significantly larger in *M. soricoides* compared to both clades of *M.'fotsifotsy* (Fig. 4.7, Supplementary Fig. 4.4), whereas *Microgale nasoloi* was generally intermediate. The north clade of *M.'fotsifotsy* was significantly smaller than the widespread clade in eight measurements (Fig. 4.7). Most measurement indices, which were used to infer ecological adaptations to arboreality, fossoriality, or terrestriality, were also significantly different between species and clades (Supplementary Fig. 4.4). However, these indices did not consistently point to an ecological adaptation in any one species; for example, a high scapular shape index suggests adaptations for climbing in *M.'fotsifotsy*, while a high intermembral index suggests adaptions for climbing in *M. soricoides*.

4.5 Discussion

This study highlights the continuous nature of the speciation process with examples of speciation with and without ongoing gene flow, geographic overlap, and morphological divergence. First, we found that *Microgale soricoides* and the widespread clade of *M.'fotsifotsy*, which are broadly sympatric, have experienced low levels of gene flow. Despite this incomplete reproductive isolation, both genetic and morphometric evidence continue to support the existing taxonomy. In terms of morphology, the original description of *M.'fotsifotsy* noted that these two species are proportionally similar, but "the skull of *M.'fotsifotsy* is much smaller and more delicate than that of *M. soricoides*" (Jenkins *et al.* 1997, p. 5). This was corroborated by the morphometric data in this study, which showed that *M.'fotsifotsy* was significantly smaller in all measurements.

Our top-ranking PHRAPL model includes gene flow between *M. soricoides* and *M.'fotsifotsy* during the early stages of divergence (speciation-with-gene-flow), as well as recent gene flow between *M. soricoides* and the sympatric widespread clade of *M.'fotsifotsy*. Hybridization appears to be a rare occurrence, however, as the amount of gene flow that we detected is very small (the estimated fraction of migrants in each population per generation is 0.0004) and we did not recover any substantial admixture in our fastStructure analysis. Nonetheless, the presence of gene flow between *M. soricoides* and the

widespread clade of *M. fotsifotsy* offers the most likely explanation as to why these were recovered as sister groups in our mitochondrial phylogeny. Similar patterns of mito-nuclear discord in other taxa have been linked to mitochondrial capture, i.e., the complete replacement of the mitochondrial genome of one species with that of another species due to introgression, despite negligible amounts of modern-day gene flow (Good *et al.* 2015; Hailer *et al.* 2012). Importantly, because we were not able to include *M. nasoloi* in our gene flow analyses, we cannot be certain that introgression has not occurred between this species and either of the others; however, this seems unlikely given *M. nasoloi*'s markedly different geographic range and habitat preference.

Previous studies have hypothesized that allopatric speciation in Madagascar's humid-forest taxa was driven by Quaternary paleoclimatic conditions that caused isolation of lineages (incipient species) in refugia (Boumans *et al.* 2007; Everson *et al.* 2018; Wilmé *et al.* 2006; Wollenberg *et al.* 2008). If species expanded from refugial populations, we would expect to see genetic signatures of a bottleneck and subsequent expansion (Hewitt 2000; Keppel *et al.* 2012). Our Bayesian Skyline Plots were largely inconclusive due to extremely large confidence intervals, but they do show a general pattern of increasing population sizes in the recent past. Further studies, preferably that take paleoclimate models into account, will be needed to validate the idea that the genetic signature observed in this study are related to expansion from past refugia.

Ecological divergence may have also played a role in the speciation of the *M. soricoides - M. fotsifotsy - M. nasoloi* clade. The vast majority of morphological variance among species was related to body size differences, with *M. soricoides* having a head-body length 22% larger than *M. fotsifotsy* on average, and *M. nasoloi* being intermediate in most measurements. A future study of diet would be needed to test the idea that differences in body size correspond to different food sources. The morphological indices we evaluated were also significantly different among species, but the ecological narratives that we inferred from each index conflicted with one another. For example, the humerofemoral index suggests that *M. soricoides* might be better adapted to climbing than the other species, whereas scapula shape suggests that *M. fotsifotsy* and *M. nasoloi* are better climbers (Supplementary Fig. 4.4). Importantly, all three species are most commonly collected using pitfall traps and are occasionally collected using live traps placed up to 1.5 meters off the ground; thus, it is likely that all three species are primarily terrestrial but with some ability to move on low-lying vines and branches.

With respect to the two deeply divergent clades of *M. fotsifotsy*, several lines of evidence support their recognition as distinct species. First, although the top-ranked demographic model selected in our PHRAPL analysis included gene flow between the two clades of *M. fotsifotsy*, the amount of gene flow recovered was very small, and our fastStructure analysis did not recover any genetic admixture. Furthermore, *M. fotsifotsy* (*sensu lato*) appears to be paraphyletic, as all of our data point to *M. nasoloi*

being sister to the north clade. Although the two clades of *M. fotsifotsy* could not be easily distinguished in our PCAs and CVAs, we identified eight individual measurements that are significantly different (Fig. 4.7). The measurement with the least amount of overlap between clades is tail length; the north clade's tail length ranges from 70-80 mm, while the widespread clade's tail is significantly larger, ranging from 78-93 mm. This does not appear to be the result of a latitudinal, clinal trend; a regression of tail length against latitude in *M. fotsifotsy* shows a marked shift from the widespread clade to the north clade (Supplementary Fig. 4.5). In the original description of *M. fotsifotsy*, Jenkins et al. (1997) also noted that specimens from the Parc National (PN) de la Montagne d'Ambre (representing the north clade) are "generally paler" than those from all other localities (p. 4). The holotype of *M. fotsifotsy* was collected from the PN Montagne d'Ambre, which is a north clade locality; thus, the widespread clade will be described in the peer-reviewed literature as a new species.

Although the two clades (or species) of *M. fotsifotsy* appear to be common in their respective ranges, as they are regularly collected during field surveys, their conservation statuses will need to be reassessed in light of our imminent taxonomic changes. Of particular concern is the north clade of *M. fotsifotsy*, which is known from only 2 localities (Parc National de Montagne d'Ambre and Forêt de Binara) and has lower levels of genetic diversity and estimated population sizes compared to the other two species. Visual comparison of forest cover maps from the 1950s to today shows that deforestation has occurred in the far north (Harper *et al.* 2007) although the actual extent of forest loss in this region has not been measured. According to the International Union for the Conservation of Nature's (IUCN's) RedList criteria, species that are known to exist at no more than five locations and whose habitat is inferred to be declining should be considered Endangered. Pending formal species description, we recommend that the north clade of *M. fotsifotsy* be treated as Endangered until a formal IUCN assessment of this species is conducted. Relevant to conservation, the presence of a newly recognized species endemic to the far north of Madagascar supports the emerging idea that the region surrounding Montagne d'Ambre is important for harboring and/or generating cryptic endemism (see also Chapter 3) and should continue to be recognized as a protected area (Louis *et al.* 2008; Raxworthy & Nussbaum 1994).

Finally, this study speaks to the power and importance of integrating multiple types of data. Our mtDNA and nuDNA sequence data produced different phylogenies, but using both datasets alongside morphological analyses painted a clear picture of the evolutionary history of this species complex. Such work relies on regular biological surveys, and on voucher specimens being appropriately preserved and curated in natural history collections. Without the combined efforts of field collectors and museum preparators, the patterns of speciation and gene flow, and the new species of shrew tenrec identified here, would have gone unrecognized.

4.6 Acknowledgements

This research was supported by grant DEB-1120904 (L.E.O.) and a Graduate Research Fellowship (K.M.E.) from the National Science Foundation; a Graduate Student Research Award from the Society of Systematic Biologists; a Grant-in-Aid of Research from the American Society of Mammalogists; and a Dissertation Completion Fellowship from the University of Alaska Fairbanks Graduate School. We are grateful for the many years of fieldwork conducted by the students and staff of the Vahatra Association in Madagascar, made possible through permits for collection and export granted by Madagascar National Parks and the Ministère des Forêts et de l'Environement. We applaud the excellent specimen preparation and curation by the following individuals: Christopher J. Raxworthy, Nancy B. Simmons, and Neil P. Duncan (American Museum of Natural History); Bruce D. Patterson, Lawrence R. Heaney, Adam W. Ferguson, and the late William T. Stanley (Field Museum of Natural History); Paula D. Jenkins, Louise Tomsett, Roberto Portela Miguez, and Richard Sabin (The Natural History Museum, London [formerly the British Museum of Natural History]); Daniel Rakotondravony, Steven M. Goodman, and Voahangy Soarimalala (Université d'Antananarivo, Département de Biologie Animale); Linda K. Gordon, Darrin P. Lunde, and Michael D. Carleton (National Museum of Natural History, Smithsonian); Priscilla K. Tucker, Cody W. Thompson, Steve Hinshaw, and Philip Myers (University of Michigan Museum of Zoology). For their assistance generating and analyzing the UCE data, we thank Travis C. Glenn, Brant C. Faircloth, Michael E. Alfaro, Janet C. Buckner, Kyndall B.P. Hildebrandt, Jessica F. McLaughlin, and Kevin Winker. Finally, we are grateful to the following individuals for their assistance with writing and editing: Nicholas J. Kerhoulas, Derek S. Sikes, Naoki Takebayashi, and Sharon A. Jansa.

4.7 References

Andermann T, Fernandes AM, Olsson U, Töpel M, Pfeil B, Oxelman B, Aleixo A, Faircloth BC, Antonelli A (2018) Allele phasing greatly improves the phylogenetic utility of ultraconserved elements. *Systematic Biology* (in press).

Bickford D, Lohman D, Sodhi N (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, *22*, 148–155.

Blaimer BB, Fisher BL (2013) How much variation can one ant species hold? Species delimitation in the *Crematogaster kelleri*-group in Madagascar. *PLoS ONE*, *8*, e68082.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120.

Bouckaert R, Heled J, Kühnert D *et al.* (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *10*, e1003537.

Boumans L, Vieites DR, Glaw F, Vences M (2007) Geographical patterns of deep mitochondrial differentiation in widespread Malagasy reptiles. *Molecular Phylogenetics and Evolution*, *45*, 822–839.

Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A (2012) Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, *29*, 1917–1932.

Coyne JA, Orr HA (2004) *Speciation*. Sunderland, Massachusetts: Sinauer Associates.

Drummond A, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, *7*, 214.

Eckert AJ, Carstens BC (2008) Does gene flow destroy phylogenetic signal? The performance of three methods for estimating species phylogenies in the presence of gene flow. *Molecular Phylogenetics and Evolution*, *49*, 832–842.

Eisenberg JF, Gould E (1970) The tenrecs: a study in mammalian behavior and evolution. *Smithsonian Contributions to Zoology*, *27*, 1–137.

Everson KM, Hildebrandt KBP, Goodman SM, Olson LE (2018) Caught in the act: incipient speciation across a latitudinal gradient in a semifossorial mammal from Madagascar, the mole tenrec *Oryzorictes hova* (Tenrecidae). *Molecular Phylogenetics and Evolution, 126,* 74–84.

Everson KM, Soarimalala V, Goodman SM, Olson LE (2016) Multiple loci and complete taxonomic sampling resolve the phylogeny and biogeographic history of tenrecs (Mammalia: Tenrecidae) and reveal higher speciation rates in Madagascar's humid forests. *Systematic Biology*, *65*, 890–909.

Faircloth BC (2013) illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming. http://dx.doi.org/10.6079/J9ILL.

Faircloth BC (2016) PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics*, *32*, 786–788.

Faircloth BC, Glenn TC (2012) Not all sequence tags are created equal: designing and validating sequence identification tags robust to indels. *PLoS ONE*, *7*, e42543.

Faircloth BC, McCormack JE, Crawford NG *et al.* (2012) Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, *61*, 717–726.

Good JM, Vanderpool D, Keeble S, Bi K (2015) Negligible nuclear introgression despite complete mitochondrial capture between two species of chipmunks. *Evolution*, *69*, 1961–1972.

Goodman SM, Jenkins PD (2000) Tenrecs (Lipotyphla: Tenrecidae) of the Parc National de Marojejy, Madagascar. *Fieldiana Zoology (New Series), 97*, 201–229.

Goodman SM, Raxworthy CJ, Maminirina CP, Olson LE (2006) A new species of shrew tenrec (*Microgale jobihely*) from northern Madagascar. *Journal of Zoology*, *270*, 384–398.

Goodman SM, Soarimalala V (2004) A new species of *Microgale* (Lipotyphla: Tenrecidae: Oryzorictinae) from the Forêt des Mikea of southwestern Madagascar. *Proceedings of the Biological Society of Washington*, *117*, 251–265.

Grabherr MG, Haas BJ, Yassour M *et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*, 644–652

Grinand C, Rakotomalala F, Gond V *et al.* (2013) Estimating deforestation in tropical humid and dry forests in Madagascar from 2000 to 2010 using multi-date Landsat satellite images and the random forests classifier. *Remote Sensing of Environment*, *139*, 68–80.

Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009) Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics*, *5*, e1000695.

Hailer F, Kutschera VE, Hallström BM *et al.* (2012) Nuclear genomic sequences reveal that polar bears are an old and distinct bear lineage. *Science*, *336*, 344–347.

Harper G, Steininger M, Tucker C, Juhn D, Hawkins F (2007) Fifty years of deforestation and forest fragmentation in Madagascar. *Environmental Conservation*, *34*, 325–333.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, *405*, 907–913.

Hey J (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution*, *27*, 905–920.

Huelsenbeck J, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, *17*, 754–755.

Jackson ND, Carstens BC, Morales AE, O'Meara BC (2016) Species delimitation with gene flow. *Systematic Biology*, *66*, 799-812.

Jackson ND, Morales AE, Carstens BC, O'Meara BC (2017) PHRAPL: phylogeographic inference using approximate likelihoods. *Systematic Biology*, *66*, 1045–1053.

Jenkins PD (1993) A new species of *Microgale* (Insectivora, Tenrecidae) from eastern Madagascar with an unusual dentition. *American Museum Novitates*, *3067*, 1–11.

Jenkins PD (2003) *Microgale*, shrew tenrecs. In S. Goodman & J. Benstead (Eds.), *The Natural History of Madagascar* (pp. 1273–1278). University Of Chicago Press.

Jenkins PD, Goodman SM (1999) A new species of *Microgale* (Lipotyphla, Tenrecidae) from isolated forest in southwestern Madagascar. *Bulletin of the Natural History Museum of London (Zoology)*, *65*, 155–164.

Jenkins PD, Raxworthy CJ, Nussbaum RA (1997) A new species of *Microgale* (Insectivora, Tenrecidae), with comments on the status of four other taxa of shrew tenrecs. *Bulletin of the Natural History Museum of London (Zoology)*, *63*, 1–96.

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, *30*, 772–780.

Keppel G, Van Niel KP, Wardell-Johnson GW *et al.* (2012) Refugia: identifying and understanding safe havens for biodiversity under climate change. *Global Ecology and Biogeography*, *21*, 393–404.

Leaché AD, Harris RB, Rannala B, Yang Z (2014) The influence of gene flow on species tree estimation: a simulation study. *Systematic Biology*, *63*, 17–30.

Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, *27*, 2987–2993.

Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint*, 1303.3997.

Louis EE, Engberg SE, McGuire SM *et al.* (2008) Revision of the mouse lemurs, *Microcebus* (Primates, Lemuriformes), of northern and northwestern Madagascar with descriptions of two new species at Montagne d'Ambre National Park and Antafondro Classified Forest. *Primate Conservation*, *23*, 19–38.

Miralles A, Vences M (2013) New metrics for comparison of taxonomies reveal striking discrepancies among species delimitation methods in *Madascincus* lizards. *PLoS ONE*, *8*, e68242.

Mirarab S, Reaz R, Bayzid MS *et al.* (2014) ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics*, *30*, 541–548.

Morales AE, Jackson ND, Dewey TA, O'Meara BC, Carstens BC (2016) Speciation with gene flow in North American *Myotis* bats. *Systematic Biology*, *66*, 440-452.

Nosil P (2008) Speciation with gene flow could be common. *Molecular Ecology*, *17*, 2103–2106.

Olivieria G, Zimmermann E, Randrianambinina B *et al.* (2007) The ever-increasing diversity in mouse lemurs: three new species in north and northwestern Madagascar. *Molecular Phylogenetics and Evolution*, *43*, 309–327.

Olson LE, Goodman SM, Yoder AD (2004) Illumination of cryptic species boundaries in long-tailed shrew tenrecs (Mammalia: Tenrecidae; *Microgale*), with new insights into geographic variation and distributional constraint. *Biological Journal of the Linnean Society*, *83*, 1–22.

Olson LE, Rakotomalala Z, Hildebrandt KBP *et al.* (2009) Phylogeography of *Microgale brevicaudata* (Tenrecidae) and description of a new species from western Madagascar. *Journal of Mammalogy*, *90*, 1095–1110.

Pinho C, Hey J (2010) Divergence with gene flow: models and data. *Annual Review of Ecology, Evolution, and Systematics*, *41*, 215–230.

Poux C, Madsen O, Glos J, Jong W de, Vences M (2008) Molecular phylogeny and divergence times of Malagasy tenrecs: influence of data partitioning and taxon sampling on dating analyses. *BMC*

Evolutionary Biology, *8*, 102–118.

R Core Team (2012) R: a language and environment for statistical computing. Available from http://www.R-project.org.

Raj A, Stephens M, Pritchard JK (2014) fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*, *197*, 573–589.

Rambaut A, Drummond A (2007) Tracer v1.6. Available from http://beast.community/tracer.

Raxworthy CJ, Nussbaum RA (1994) A rainforest survey of amphibians, reptiles and small mammals at Montagne d'Ambre, Madagascar. *Biological Conservation*, *69*, 65–73.

Rosenberg NA (2003) distruct: a program for the graphical display of population structure. *Molecular Ecology Notes*, *4*, 137–138.

Sáez AG, Lozano E (2005) Body doubles. *Nature*, *433*, 111.

Salton JA, Sargis EJ (2008) Evolutionary morphology of the Tenrecoidea (Mammalia)forelimb skeleton. In E. J. Sargis & M. Dagosto (Eds.), *Mammalian Evolutionary Morphology* (pp. 51–71). Dordrecht, Netherlands: Springer.

Salton JA, Sargis EJ (2009) Evolutionary morphology of the Tenrecoidea (Mammalia) hindlimb skeleton. *Journal of Morphology*, *270*, 367–387.

Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*, 1312–1313.

Stange M, Sánchez-Villagra MR, Salzburger W, Matschiner M (2018) Bayesian divergence-time estimation with genome-wide single-nucleotide polymorphism data of sea catfishes (Ariidae) supports Miocene closure of the Panamanian isthmus. *Systematic Biology, 67,* 681-699.

van Buuren S, Groothuis-Oudshoorn K (2010) MICE: multivariate imputation by chained equations in R. *Journal of Statistical Software*, 1–68.

Vences M, Wollenberg KC, Vieites DR, Lees DC (2009) Madagascar as a model region of species diversification. *Trends in Ecology & Evolution*, *24*, 456–465.

Vieites DR, Wollenberg KC, Andreone F *et al.* (2009) Vast underestimation of Madagascar's biodiversity evidenced by an integrative amphibian inventory. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 8267–8272.

Wickham H (2016) *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Wilmé L, Goodman SM, Ganzhorn JU (2006) Biogeographic evolution of Madagascar's microendemic biota. *Science*, *312*, 1063–1065.

Wollenberg KC, Vieites DR, van der Meijden A *et al.* (2008) Patterns of endemism and species richness in Malagasy cophyline frogs support a key role of mountainous areas for speciation. *Evolution*, *62*, 1890–1907.

Woodman N, Stabile FA (2015) Functional skeletal morphology and its implications for locomotory behavior among three genera of myosoricine shrews (Mammalia: Eulipotyphla: Soricidae). *Journal of Morphology*, *276*, 550–563.

Xi Z, Liu L, Davis CC (2015) Genes with minimal phylogenetic information are problematic for coalescent analyses when gene tree estimation is biased. *Molecular Phylogenetics and Evolution*, *92*, 63–71.

4.8 Tables

Table 4.1. Descriptions of all morphological measurements and indices. All cranioskeletal measurements were taken as minimum distance between landmarks. Dental loci are identified by tooth type (I/i, upper/lower incisor; C/c, upper/lower canine; P/p, upper/lower premolar; and M/m, upper/lower molar) and position, e.g. I3 refers to the phylogenetic third upper incisor of the permanent dentition. We follow the dental nomenclature of MacPhee (1987). Table continues across two pages.

Abbreviation	Module	Measurement description
BB	Cranium	Braincase breadth: Greatest cranial breadth, as measured across squamosals.
C1W	Cranium	Greatest width across C1.
CC1	Cranium	Condylo-C1 length: Caudal surface of occipital condyle to rostral surface of C1.
CEG	Cranium	Condylo-entoglenoid length: Posterior-most (caudal) surface of occipital condyle to anterior-most (rostral) surface of the entoglenoid process of the squamosal.
CI2	Cranium	Condylo-I2 length: Posterior-most (caudal) surface of occipital condyle to anterior-most (rostral) surface of I2.
CIL	Cranium	Condyloincisive length: Posterior-most (caudal) surface of occipital condyle to anterior-most (rostral) surface of I1.
CPM	Cranium	Condylopremaxillary length: Posterior-most (caudal) surface of occipital condyle to anterior-most (rostral) surface of the premaxilla.
I2W	Cranium	Greatest width across I2.
M1WA	Cranium	Greatest anterior width across M1: Greatest breadth across M1 as measured from lateral surface of anterior ectostyle and/or mesiostyle.
M1WP	Cranium	Greatest posterior width across M1: Greatest breadth across M1 as measured from lateral surface of distostyle.
M2WA	Cranium	Greatest anterior width across M2: Greatest breadth across M2 as measured from lateral surface of anterior ectostyle and/or mesiostyle.
M2WP	Cranium	Greatest posterior width across M2: Greatest breadth across M2 as measured from labial surface of distostyle.
M3W	Cranium	Greatest breadth across M3: Greatest breadth across M3, as measured from labial surface of distostyle.
OP	Cranium	Posterior-most (caudal) surface of the occipital to anterior-most (rostral) surface of the premaxilla.
P2W	Cranium	Greatest width across P2.
P3W	Cranium	Greatest width across P3.
P4M3	Cranium	Greatest distance between P4 and M3 as measured from the rostral surface of P4 to the caudal surface of M3.
P4WP	Cranium	Greatest posterior width across P4: Greatest breadth across P4 as measured from labial surface of distostyle.
PPMX	Cranium	Paroccipital process to premax length: Caudal surface of paroccipital process of basioccipital to anteriormost (rostral) surface of the premaxilla.
PEG	Cranium	Paroccipital process to entoglenoid length: Rostral surface of entoglenoid process to caudal surface of paroccipital process of basioccipital.
PZ	Cranium	Premaxillary to zygomatic length: Rostral surface of premaxilla to caudal surface of zygomatic process of maxilla.
UTR	Cranium	Upper toothrow length: Rostral surface of I1 to caudal surface of M3.
ZB	Cranium	Zygomatic breadth: Greatest breadth across maxillary zygomatic processes.
AC	Mandible	Angular process depth: Greatest distance between angular and condyloid processes of mandible.
Ai1	Mandible	Greatest length of mandible: Greatest distance between caudal surface of angular process of mandible and rostral surface of i1.
CCor	Mandible	Height of coronoid process: Greatest distance between ventral surface of condyloid process and dorsal surface of coronoid process.
Ci1	Mandible	Greatest distance between the the coronoid process and i1.

m3c1	Mandible	Greatest distance between m3 and c1 as measured from the rostral surface of c1 to caudal surface of m3.
m3i1	Mandible	Greatest distance between m3 and i1 as measured from the rostral surface of c1 to caudal surface of i1.
MCW	Mandible	Mandibular condyle width: Greatest breadth across buccal and labial surfaces of mandibular condyle.
MH	Mandible	Height of mandible: Greatest distance between coronoid and angular processes of mandible.
ML	Mandible	Condylo-i1 length: Greatest distance between caudal surface of condyloid process of mandible and rostral surface of i1. ML2 of Olson et al. (2004).
FGT	Postcranial skeleton	Femur, lateral condyle to greater trochanter (FGT): Distal surface of lateral condyle to proximal surface of greater trochanter.
FMH	Postcranial skeleton	Femur, medial condyle to head (FMH): Distal surface of medial condyle to proximal surface of femoral head.
TIB	Postcranial skeleton	Tibia, greatest length (TIB).
HCW	Postcranial skeleton	Humerus, distal width (HCW): Greatest breadth across distal condyles.
HHC	Postcranial skeleton	Humerus, head to capitulum length (HHC): Proximal surface of humeral head to distal surface of capitulum.
HHME	Postcranial skeleton	Humerus, head to medial epicondyle (HHME): Proximal surface of humeral head to distal surface of medial epicondyle.
HGL	Postcranial skeleton	Humerus, greatest length (HGL): Proximal surface of humeral head to distal surface of trochlear.
RAD	Postcranial skeleton	Radius, greatest length (RAD).
ULN	Postcranial skeleton	Ulna, greatest length (ULN).
ScL	Postcranial skeleton	Scapular length (ScL): Greatest length of scapula from coracoid process to inferior angle.
ScW	Postcranial skeleton	Scapular width (ScW): Width of scapula as measured by placing one caliper jaw against both the lateral margin of glenoid cavity and axillary border and the other against the superior border.
PelvisGrLngth	Postcranial skeleton	Greatest length of the pelvis.
IW	Postcranial skeleton	Pelvis, greatest width across anterior ilia (IW).
HB	External	Head and body length: Tip of nose and caudal point of the body (at base of tail). Counterintuitively, this is often unequal to the difference between TL and TV.
TV	External	Tail length: external length of tail.
HF	External	Length of hindfoot including claws.
Relative scapular width	N/A (index)	Ratio of scapular width (ScW) to scapular length (ScL). Short, broad scapulae are associated with climbing, while long and narrow scapulae are associated with terrestriality/running (Salton & Sargis 2008)
Brachial Index	N/A (index)	Ratio of greatest radius length to greatest humerus length (RAD/HGL). This index decreases with increasing fossoriality in rodents (Samuels & Van Valkenburgh 2008; Wooman & Stabile 2015).
Crural Index	N/A (index)	Ratio of greatest tibia length to greatest femur length (TIB/FGT). Lower values are associated with fossoriality in tenrecs (Salton & Sargis 2009).
Humerofemoral Index	N/A (index)	Ratio of greatest humerus length to greatest femur length (HGL/FGT). A high index is associated with climbing in some small mammals (Sargis 2002).
HCW:HGL	N/A (index)	Ratio of distal humerus width (HCW) to greatest humerus length (HGL). A long, narrow humerus is associated with climbing, whereas a robust, short, wide humerus is associated with digging (Salton & Sargis 2009).
Intermembral Index	N/A (index)	Ratio of the length of the forelimb (FGT+TIB) to the length of the hindlimb (HGL+RAD). This index is often larger in fossorial rodents (Samuels & Van Valkenburgh 2008; Woodman & Stabile 2015)
TV.HB	N/A (index)	Ratio of tail vertebrate length (TV) to head-body length (HB). Long tails are often associated with climbing in small mammals (Cartmill 1985).
ULN:RAD	N/A (index)	Ratio of greatest ulna length to greatest radius length (ULN/RAD). In tenrecs, higher relative ulna lengths are generally due to an elongated olecronon processare, which is associated with fossoriality (Hildebrand 1985; Woodman & Stabile 2015)

Table 4.2. Classification matrices from our canonical variates analyses, showing the number of individuals from each clade included in the analysis (rows) and the groups to which they were assigned in the analysis (columns).

Cranial Measurements Only			
	M. fotsifotsy (north)	*M. fotsifotsy* (widespread)	*M. soricoides*
M. fotsifotsy (north)	5	3	0
M. fotsifotsy (widespread)	0	38	0
M. soricoides	0	0	79
Mandibular Measurements Only			
	M. fotsifotsy (north)	*M. fotsifotsy* (widespread)	*M. soricoides*
M. fotsifotsy (north)	0	8	0
M. fotsifotsy (widespread)	0	38	0
M. soricoides	0	1	78
Postcranial Measurements Only			
	M. fotsifotsy (north)	*M. fotsifotsy* (widespread)	*M. soricoides*
M. fotsifotsy (north)	0	2	0
M. fotsifotsy (widespread)	0	20	0
M. soricoides	0	0	49
External/Field Measurements Only			
	M. fotsifotsy (north)	*M. fotsifotsy* (widespread)	*M. soricoides*
M. fotsifotsy (north)	3	1	0
M. fotsifotsy (widespread)	0	28	1
M. soricoides	0	0	63
All Measurements			
	M. fotsifotsy (north)	*M. fotsifotsy* (widespread)	*M. soricoides*
M. fotsifotsy (north)	1	1	0
M. fotsifotsy (widespread)	0	26	0
M. soricoides	0	0	51

4.9 Figures

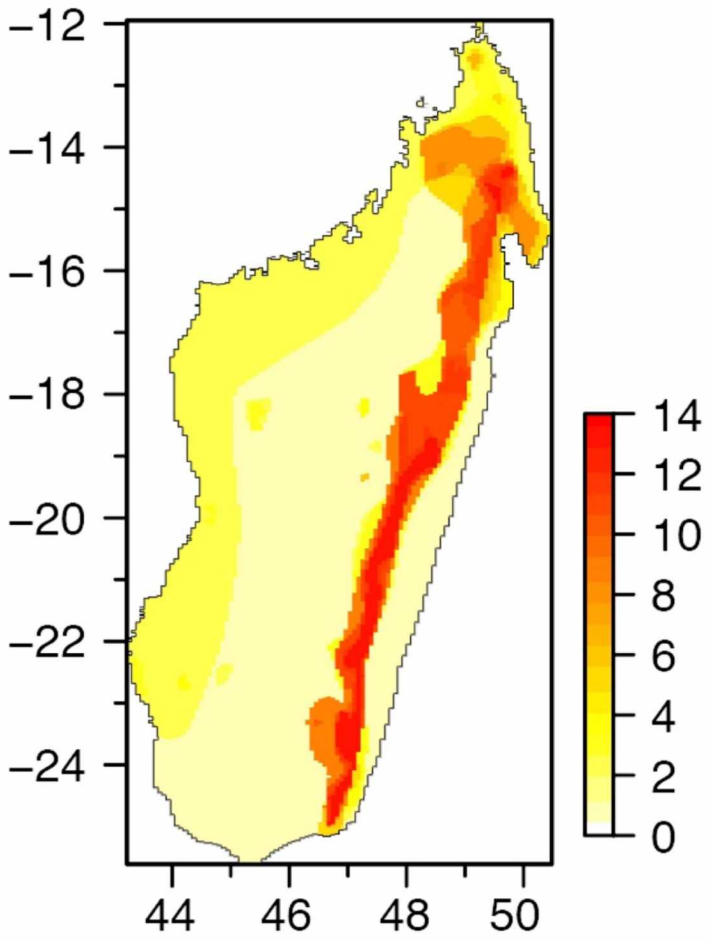


Figure 4.1. A map of Madagascar showing *Microgale* species density. Most species are broadly sympatric in the eastern humid-forest belt. Map was produced using shapefiles provided by the International Union for the Conservation of Nature (IUCN).

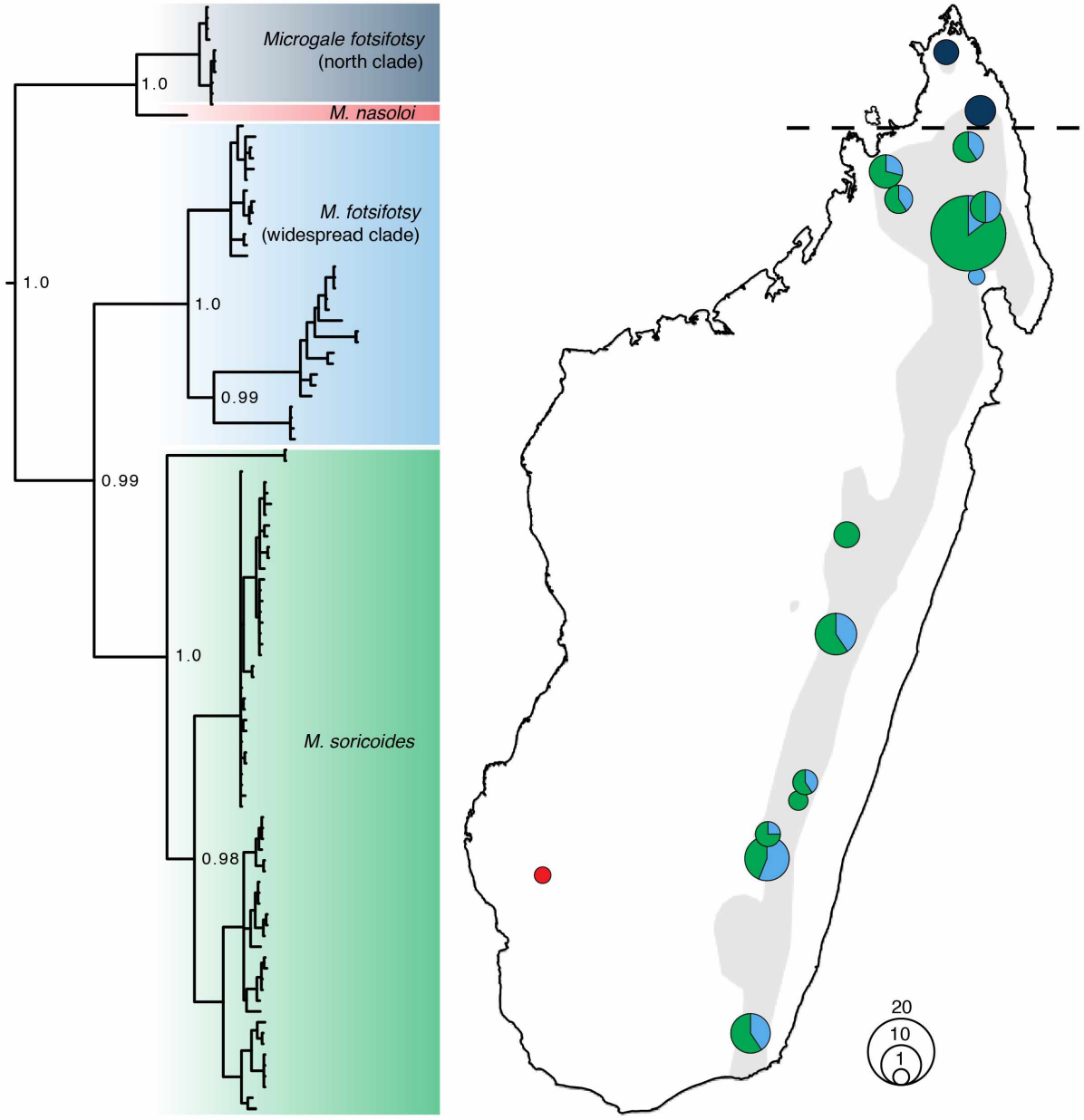


Figure 4.2. Mitochondrial phylogeny of the *Microgale fotsifotsy-soricoides-nasoloi* clade (left) and collection localities for all individuals (right). Clades and localities are colored according to species: *M. fotsifotsy* north clade (dark blue), *M. fotsifotsy* widespread clade (light blue), *M. soricoides* (green), and *M. nasoloi* (red). Pie charts show the relative proportions *M. fotsifotsy* and *M. soricoides* individuals collected from that locality. The sizes of the pie charts correspond to the number of individuals collected. Bayesian posterior probabilities are shown at the nodes of select divergence events. Phylogeny was rooted on the outgroup *M. cowani* (not shown).

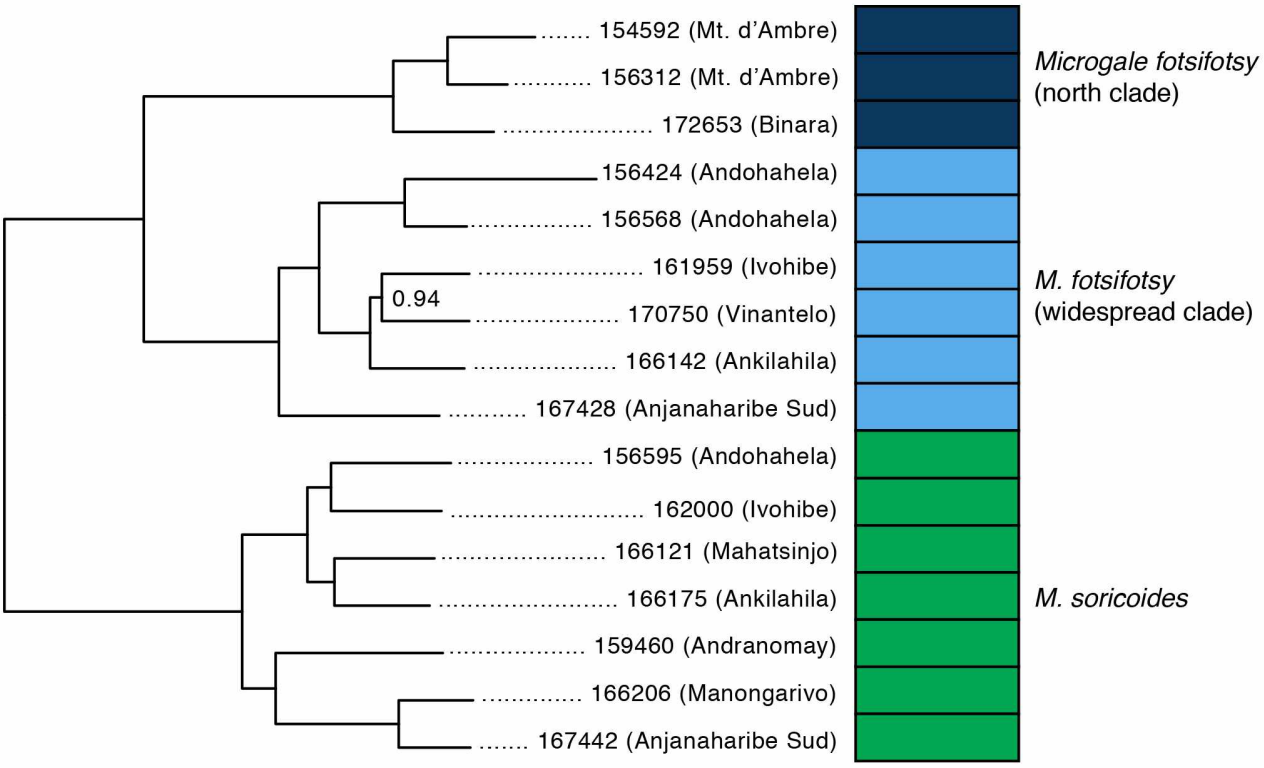


Figure 4.3. Bayesian phylogeny generated from UCE-based concatenated sequence data (left) and the corresponding admixture plots for each individual (right). Individuals are labeled with their FMNH catalog numbers, with abbreviated locality names given in parentheses. All Bayesian posterior probabilities are 1.0 unless otherwise indicated. Note that *M. nasoloi* was not included in this analysis. Phylogeny was rooted on the outgroup *M. cowani* (not shown).

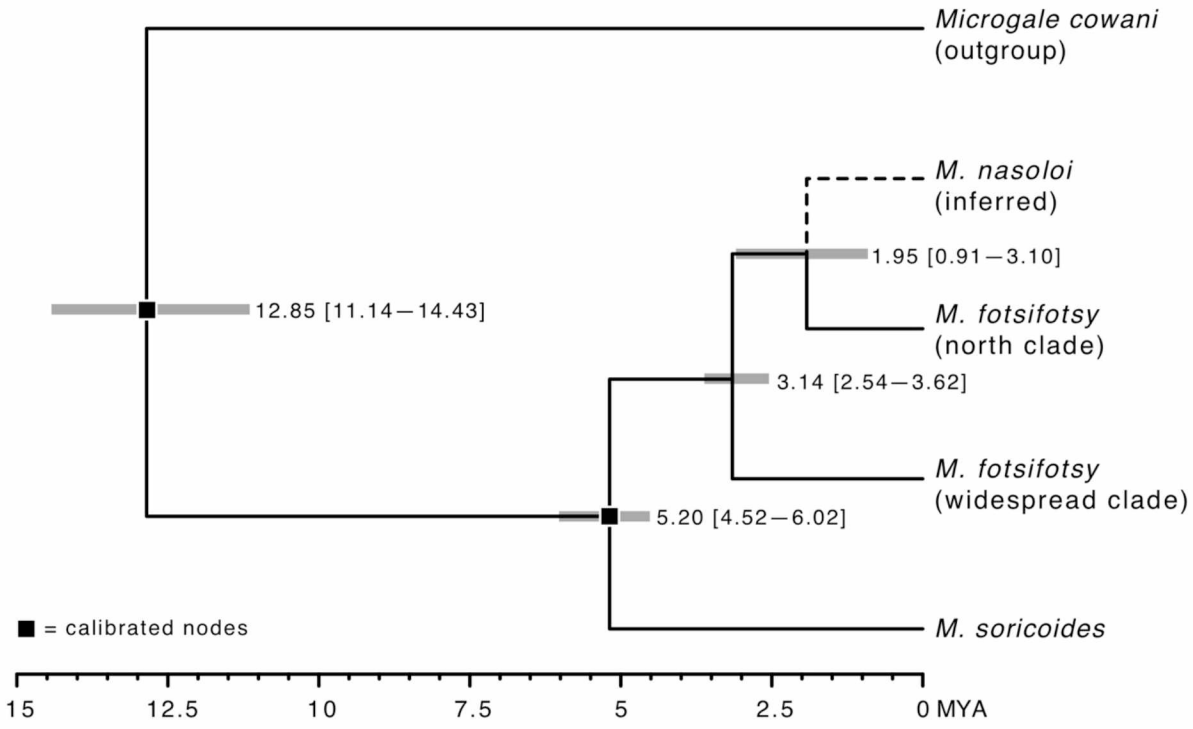


Figure 4.4. Species tree produced using the SNP dataset in SNAPP. Divergence times were estimated following procedures outlined in Stange et al. (2018) and using two calibration points from Everson et al. (2016). A summary-based species tree method (ASTRAL) produced the same topology and is therefore not shown.

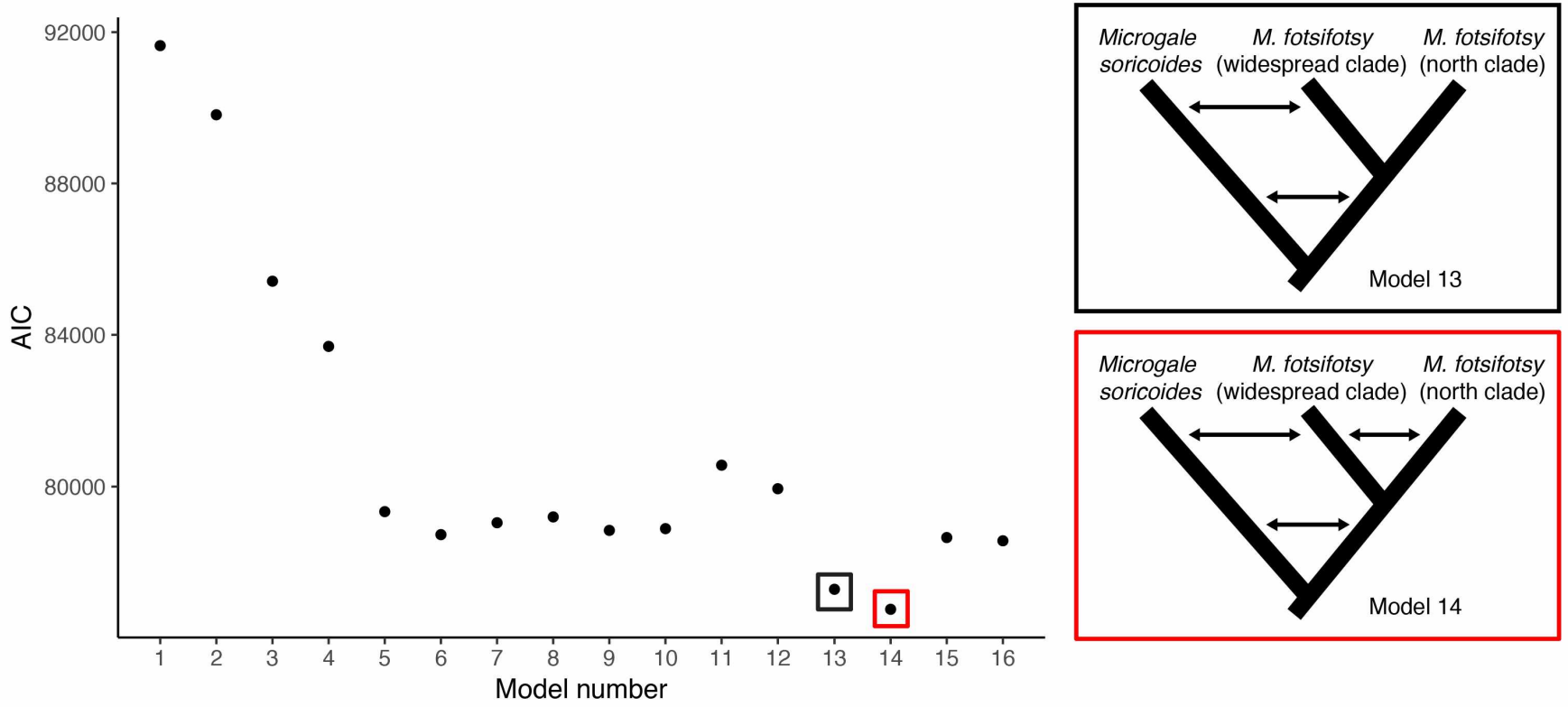


Figure 4.5. Model test results (AIC values) from our PHRAPL demographic analysis. The top two models (13 and 14) are shown to the right of the graph in black and red boxes, respectively; all other models are shown in Supplementary Fig. 4.1. Note that *Microgale nasoloi* could not be included in these analyses (see text).

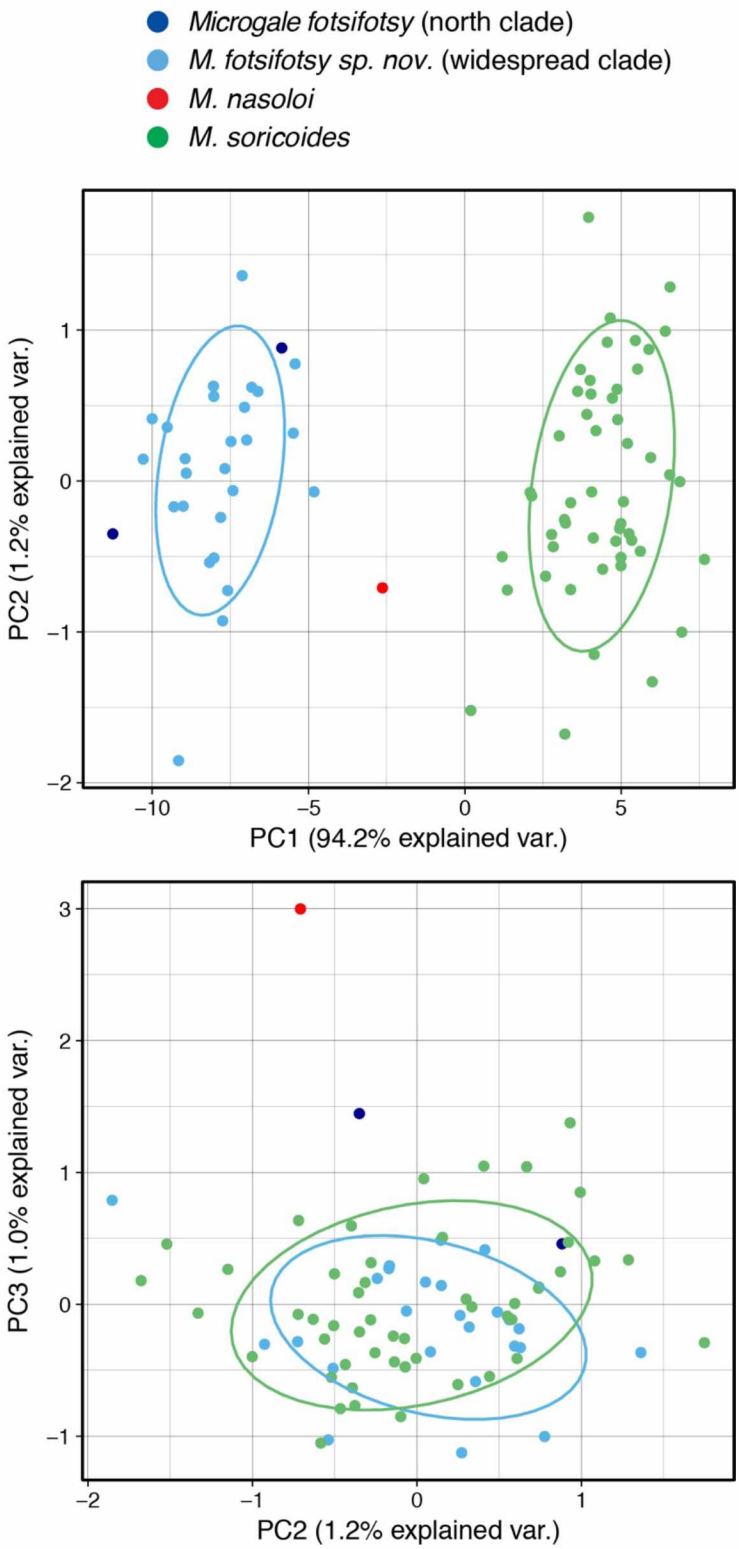


Figure 4.6. Bivariate plot of the first two principal components (a) and the second and third principal components (b) for the complete morphometric dataset, which included cranial, mandibular, postcranial, and external measurements. Individual points (and 95% confidence ellipses for clusters with >2 individuals) are colored according to species, as in Fig. 4.2.

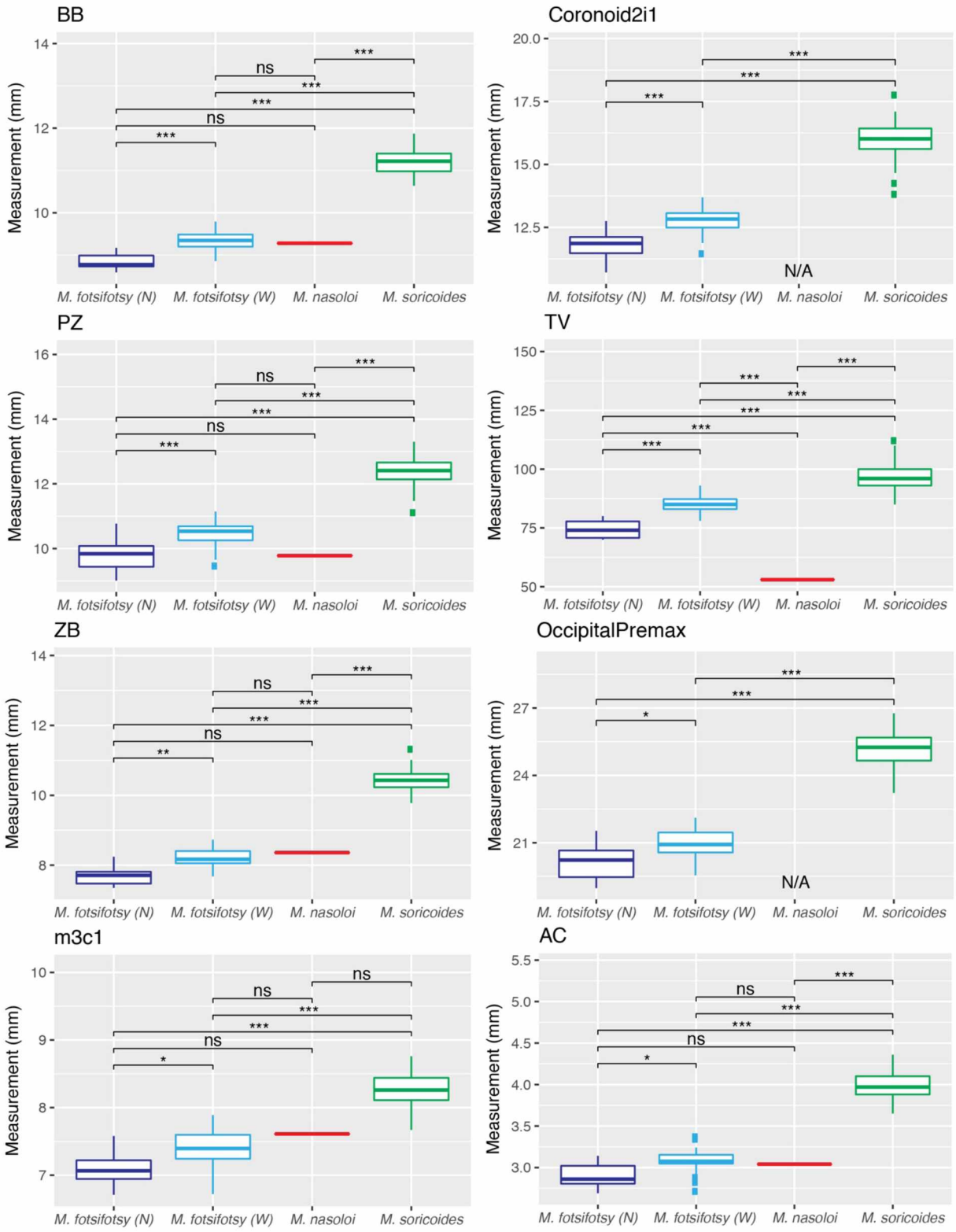


Figure 4.7. Boxplots of all morphological measurements evaluated in this study that were significantly different between the north (N) and widespread (W) clades of *Microgale fotsifotsy*. Brackets above each boxplot indicate significance values from pairwise ANOVAs. Boxes are colored according to species, as in Fig. 4.2.

4.10 Supplementary Tables

Supplementary Table 4.1. Museum catalog numbers and localities for specimens associated with sequence data. Collection abbreviations are as follows: Field Museum of Natural History (FMNH), Université d'Antananarivo Département de Biologie Animale (UADBA), American Museum of Natural History (AMNH). Table continues on multiple pages.

Collection	Catalog #	Longitude	Latitude	Locality	Locality Description
Microgale nasoloi					
FMNH	161575	44.192	-22.678	Analavelona	Forêt d'Analavelona, Antanimena, 12.5 km NW Andranoheza
Microgale fotsifotsy (north clade)					
FMNH	154591	49.172	-12.527	Montagne d'Ambre	Parc National de Montagne d'Ambre, 5.5 km SW Joffreville
FMNH	154592	49.172	-12.527	Montagne d'Ambre	Parc National de Montagne d'Ambre, 5.5 km SW Joffreville
FMNH	154593	49.172	-12.527	Montagne d'Ambre	Parc National de Montagne d'Ambre, 5.5 km SW Joffreville
FMNH	156312	49.170	-12.520	Montagne d'Ambre	Parc National de Montagne d'Ambre, 5.5 km SW Joffreville
FMNH	172653	49.617	-13.255	Binara	Forêt de Binara, 7.5 km SW Daraina
FMNH	172654	49.617	-13.255	Binara	Forêt de Binara, 7.5 km SW Daraina
UADBA	45941	49.602	-13.263	Binara	Forêt de Binara, Andohananalamazava
UADBA	45942	49.602	-13.263	Binara	Forêt de Binara, Andohananalamazava
UADBA	45957	49.602	-13.263	Binara	Forêt de Binara, Andohananalamazava
UADBA	46006	49.582	-13.240	Binara	Forêt de Binara, Antsahandrapaka
Microgale fotsifotsy (widespread clade)					
AMNH	275282	49.445	-13.700	Sorata	4 km WNW Ambodimandresy, Forêt de Sorata
AMNH	275283	49.445	-13.700	Sorata	4 km WNW Ambodimandresy, Forêt de Sorata
FMNH	156424	46.735	-24.584	Andohahela	Parc National d'Andohahela, 13.5 km NW Eminiminy, parcel 1
FMNH	156568	46.765	-24.626	Andohahela	Parc National d'Andohahela, 8 km NW Eminiminy
FMNH	156569	46.765	-24.626	Andohahela	Parc National d'Andohahela, 8 km NW Eminiminy
FMNH	156570	46.731	-24.569	Andohahela	Parc National d'Andohahela, 15 km NW Eminiminy, parcel 1
FMNH	159663	49.742	-14.437	Marojejy	Parc National de Marojejy, 11 km NW Manantenina, Antranohofa
FMNH	161815	46.960	-22.470	Ivohibe	About 7.5 km ENE Ivohibe, along Hefitany River
FMNH	161959	46.960	-22.470	Ivohibe	About 7.5 km ENE Ivohibe, along Hefitany River
FMNH	161960	46.960	-22.470	Ivohibe	About 7.5 km ENE Ivohibe, along Hefitany River
FMNH	161961	46.955	-22.497	Ivohibe	Réserve Spéciale d'Ivohibe, 6.5 km ESE Ivohibe, at source of Andranomainty River
FMNH	161965	46.938	-22.427	Ivohibe	9 km NE Ivohibe, 6.5 km ESE Angodongodona
FMNH	161967	46.938	-22.427	Ivohibe	9 km NE Ivohibe, 6.5 km ESE Angodongodona
FMNH	165776	46.946	-22.171	Andringitra	Parc National d'Andringitra, 8.5 km SE Antanifotsy
FMNH	166141	47.835	-19.707	Ankilahila	16.2 km SE Tsinjoarivo, Forêt d'Ankilahila, along Andrindrimbolo River

FMNH	166142	47.835	-19.707	Ankilahila	16.2 km SE Tsinjoarivo, Forêt d'Ankilahila, along Andrindrimbolo River
FMNH	166143	47.835	-19.707	Ankilahila	16.2 km SE Tsinjoarivo, Forêt d'Ankilahila, along Andrindrimbolo River
FMNH	166144	47.835	-19.707	Ankilahila	16.2 km SE Tsinjoarivo, Forêt d'Ankilahila, along Andrindrimbolo River
FMNH	166230	48.428	-14.000	Manongarivo	Réserve Spéciale de Manongarivo, 14.5 km SW Antanambao
FMNH	166231	48.428	-14.000	Manongarivo	Réserve Spéciale de Manongarivo, 14.5 km SW Antanambao
FMNH	167428	49.442	-14.783	Anjanaharibe-Sud	Western slope of Anjanaharibe-Sud, 13.5 km SW Befingitra
FMNH	167429	49.442	-14.783	Anjanaharibe-Sud	Western slope of Anjanaharibe-Sud, 13.5 km SW Befingitra
FMNH	170748	47.410	-21.512	Andrambovato	2 km W Andrambovato, along Tatamaly River
FMNH	170750	47.347	-21.777	Vinantelo	Forêt de Vinantelo, foot of Mont Ambodivohitra, 15.5 km SE Vohitrafeno
FMNH	172575	49.620	-14.437	Marojejy	Parc National de Marojejy, 13.0 km SE Doany
FMNH	172576	49.620	-14.437	Marojejy	Parc National de Marojejy, 13.0 km SE Doany
FMNH	176396	49.548	-15.290	Ankirindro	Makira Plateau, Forêt d'Ankirindro, 5.5 km N Marovovonana
FMNH	225993	48.586	-14.341	Bemaniveka	Province de Mahajanga; Région Sofia, Bemanivika Forest, along trail to Matsaborimena Lake
FMNH	225994	48.586	-14.341	Bemaniveka	Province de Mahajanga; Région Sofia, Bemanivika Forest, along trail to Matsaborimena Lake
SMG	11079	49.425	-14.610	Betaolana	Forêt de Betaolana, 11 km NW Ambodiangezoka
M. soricoides					
AMNH	275292	49.427	-14.531	Ambodivoara	7.43 km NW Ambodivoara
AMNH	275293	49.442	-13.686	Sorata	5 km NW Ambodimandresy, Forêt de Sorata
AMNH	275294	49.442	-13.686	Sorata	5 km NW Ambodimandresy, Forêt de Sorata
AMNH	275296	49.442	-13.686	Sorata	5 km NW Ambodimandresy, Forêt de Sorata
AMNH	275297	49.445	-13.700	Sorata	4 km WNW Ambodimandresy, Forêt de Sorata
FMNH	154029	49.462	-14.745	Anjanaharibe-Sud	Réserve Spéciale d'Anjanaharibe-Sud, 9.2 km WSW Befingitra
FMNH	154030	49.442	-14.742	Anjanaharibe-Sud	Réserve Spéciale d'Anjanaharibe-Sud, 11 km WSW Befingitra
FMNH	156462	46.735	-24.584	Andohahela	Parc National d'Andohahela, 13.5 km NW Eminiminy, parcel 1
FMNH	156483	46.735	-24.584	Andohahela	Parc National d'Andohahela, 13.5 km NW Eminiminy, parcel 1
FMNH	156594	46.738	-24.593	Andohahela	Parc National d'Andohahela, 12.5 km NW Eminiminy, parcel 1
FMNH	156595	46.735	-24.584	Andohahela	Parc National d'Andohahela, 13.5 km NW Eminiminy, parcel 1
FMNH	156596	46.731	-24.569	Andohahela	Parc National d'Andohahela, 15 km NW Eminiminy, parcel 1
FMNH	159458	47.955	-18.480	Anjozorobe	13 km SE Anjozorobe
FMNH	159459	47.955	-18.480	Anjozorobe	13 km SE Anjozorobe
FMNH	159460	47.955	-18.480	Anjozorobe	13 km SE Anjozorobe
FMNH	159461	47.955	-18.480	Anjozorobe	13 km SE Anjozorobe
FMNH	159688	49.742	-14.437	Marojejy	Parc National de Marojejy, 11 km NW Manantenina, Antranohofa
FMNH	159689	49.742	-14.437	Marojejy	Parc National de Marojejy, 11 km NW Manantenina, Antranohofa
FMNH	161997	46.968	-22.483	Ivohibe	Réserve Spéciale d'Ivohibe, 8 km E Ivohibe
FMNH	161998	46.968	-22.483	Ivohibe	Réserve Spéciale d'Ivohibe, 8 km E Ivohibe

FMNH	161999	46.968	-22.483	Ivohibe	Réserve Spéciale d'Ivohibe, 8 km E Ivohibe
FMNH	162000	46.968	-22.483	Ivohibe	Réserve Spéciale d'Ivohibe, 8 km E Ivohibe
FMNH	162002	46.955	-22.497	Ivohibe	Réserve Spéciale d'Ivohibe, 6.5 km ESE Ivohibe, at source of Andranomainty River
FMNH	165701	46.946	-22.171	Andringitra	Parc National d'Andringitra, 8.5 km SE Antanifotsy
FMNH	165787	46.946	-22.171	Andringitra	Parc National d'Andringitra, 8.5 km SE Antanifotsy
FMNH	165789	46.946	-22.171	Andringitra	Parc National d'Andringitra, 8.5 km SE Antanifotsy
FMNH	166120	47.770	-19.680	Mahatsinjo	10 km SE Tsinjoarivo, Forêt de Mahatsinjo, Andasivodihazo
FMNH	166121	47.770	-19.680	Mahatsinjo	10 km SE Tsinjoarivo, Forêt de Mahatsinjo, Andasivodihazo
FMNH	166122	47.770	-19.680	Mahatsinjo	10 km SE Tsinjoarivo, Forêt de Mahatsinjo, Andasivodihazo
FMNH	166151	47.835	-19.707	Ankilahila	16.2 km SE Tsinjoarivo, Forêt d'Ankilahila, along Andrindrimbolo River
FMNH	166175	47.835	-19.707	Ankilahila	16.2 km SE Tsinjoarivo, Forêt d'Ankilahila, along Andrindrimbolo River
FMNH	166205	48.418	-14.022	Manongarivo	Réserve Spéciale de Manongarivo, 17.3 km SW Antanambao
FMNH	166206	48.418	-14.022	Manongarivo	Réserve Spéciale de Manongarivo, 17.3 km SW Antanambao
FMNH	166207	48.418	-14.022	Manongarivo	Réserve Spéciale de Manongarivo, 17.3 km SW Antanambao
FMNH	167436	49.442	-14.783	Anjanaharibe-Sud	Western slope of Anjanaharibe-Sud, 13.5 km SW Befingitra
FMNH	167437	49.442	-14.783	Anjanaharibe-Sud	Western slope of Anjanaharibe-Sud, 13.5 km SW Befingitra
FMNH	167438	49.442	-14.783	Anjanaharibe-Sud	Western slope of Anjanaharibe-Sud, 13.5 km SW Befingitra
FMNH	167440	49.432	-14.765	Anjanaharibe-Sud	Western slope of Anjanaharibe-Sud, 13.0 km SW Befingitra
FMNH	167441	49.432	-14.765	Anjanaharibe-Sud	Western slope of Anjanaharibe-Sud, 13.0 km SW Befingitra
FMNH	167442	49.432	-14.765	Anjanaharibe-Sud	Western slope of Anjanaharibe-Sud, 13.0 km SW Befingitra
FMNH	167443	49.432	-14.765	Anjanaharibe-Sud	Western slope of Anjanaharibe-Sud, 13.0 km SW Befingitra
FMNH	167444	49.432	-14.765	Anjanaharibe-Sud	Western slope of Anjanaharibe-Sud, 13.0 km SW Befingitra
FMNH	167504	49.425	-14.610	Betaolana	Forêt de Betaolana, 11 km NW Ambodiangezoka
FMNH	167505	49.425	-14.610	Betaolana	Forêt de Betaolana, 11 km NW Ambodiangezoka
FMNH	167507	49.425	-14.610	Betaolana	Forêt de Betaolana, 11 km NW Ambodiangezoka
FMNH	167508	49.425	-14.610	Betaolana	Forêt de Betaolana, 11 km NW Ambodiangezoka
FMNH	167577	47.042	-22.163	Manambolo	Forêt de Manambolo, W slope Mont Vohipia, 19.5 km SE Sendrisoa
FMNH	167622	47.042	-22.163	Manambolo	Forêt de Manambolo, W slope Mont Vohipia, 19.5 km SE Sendrisoa
FMNH	169236	47.484	-21.584	Mandriandry	Mandriandry, 4.4 km SW Tolongoina
FMNH	170765	47.433	-21.290	Ranomafana	Parc National de Ranomafana, Vatoharanana, 4 km SW Ranomafana (ville)
FMNH	170766	47.410	-21.512	Andrambovato	2 km W Andrambovato, along Tatamaly River
FMNH	172584	49.620	-14.437	Marojejy	Parc National de Marojejy, 13.0 km SE Doany
FMNH	226014	48.586	-14.341	Bemaniveka	Province de Mahajanga; Région Sofia, Bemanivika Forest, along trail to Matsaborimena Lake
SMG	10893	48.418	-14.000	Manongarivo	Réserve Spéciale de Manongarivo, 17.3 km SW d'Antanambao, site 3
SMG	11128	49.442	-14.783	Anjanaharibe-Sud	Réserve Spéciale d'Anjanaharibe-Sud, 13.5 km SW de Befingotra, western slope, site 1
SMG	11193	49.442	-14.783	Anjanaharibe-Sud	Réserve Spéciale d'Anjanaharibe-Sud, 13.5 km SW de Befingotra, western slope, site 1

UADBA	11529	47.835	-19.707	Ankilahila	16.2 km SE Tsinjoarivo, Ankilahila along Andrindrimbolo River
UADBA	11547	48.418	-14.000	Manongarivo	Réserve Spéciale de Manongarivo, 17.3 km SW d'Antanambao, site 3
UADBA	32506	49.442	-14.783	Anjanaharibe-Sud	Réserve Spéciale d'Anjanaharibe-Sud, 13.5 km SW de Befingotra, western slope, site 1
UADBA	32507	49.442	-14.783	Anjanaharibe-Sud	Réserve Spéciale d'Anjanaharibe-Sud, 13.5 km SW de Befingotra, western slope, site 1
SMG	18319	48.586	-14.341	Bemaniveka	Province de Mahajanga; Région Sofia, Bemanivika Forest, along trail to Matsaborimena Lake
TMR	654	48.586	-14.341	Bemaniveka	Province de Mahajanga; Région Sofia, Bemanivika Forest, along trail to Matsaborimena Lake
M. cowani (outgroup)					
FMNH	161928	46.898	-22.422	Ivohibe	8 km NE Ivohibe, 5.5 km SE Angodongodona
FMNH	166102	47.770	-19.680	Mahatsinjo	10 km SE Tsinjoarivo, Forêt de Mahatsinjo, Andasivodihazo

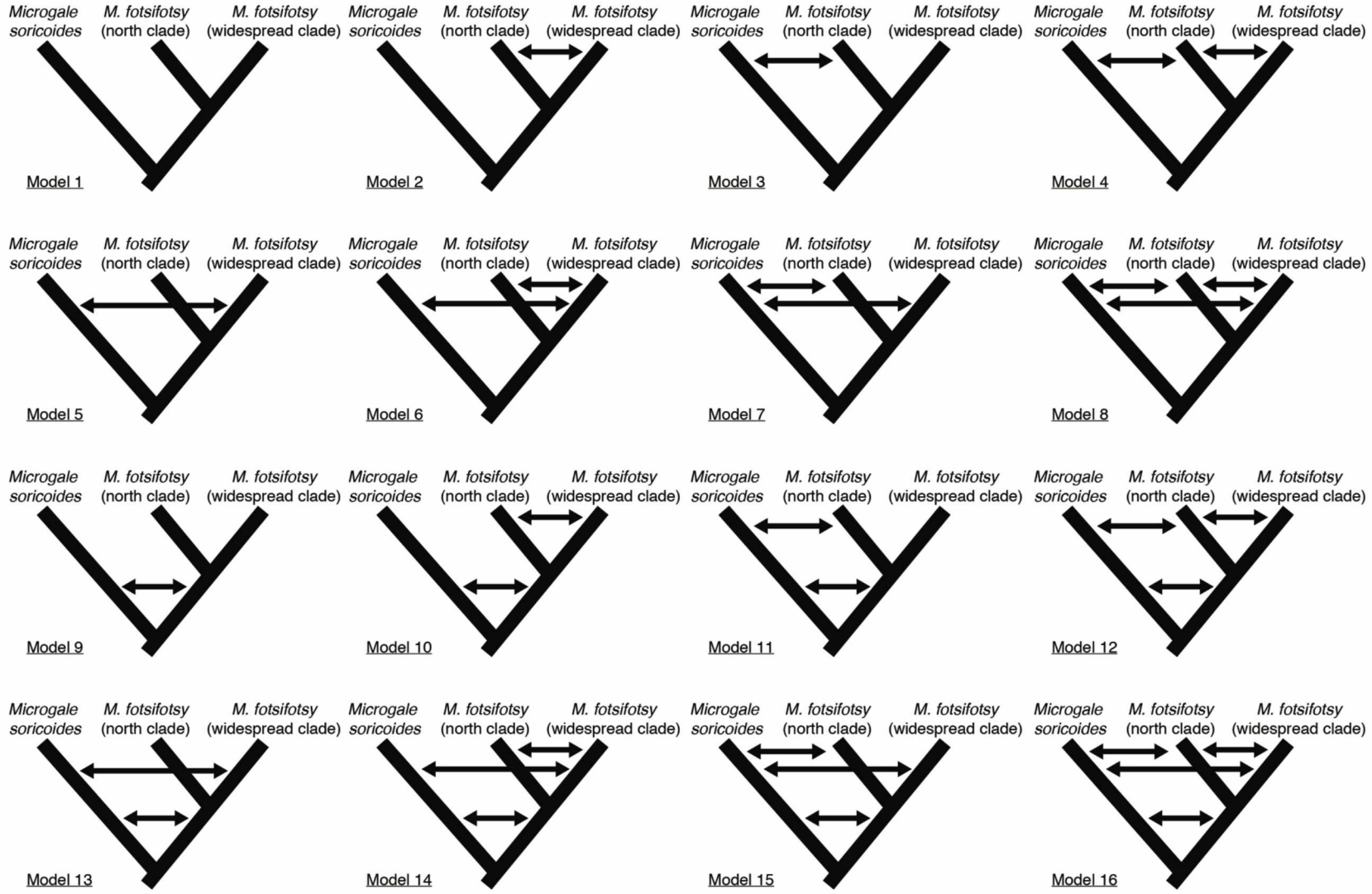
Supplementary Table 4.2. Sequencing and alignment statistics for the UCE dataset. See Supplementary Table 4.1 for information on individual localities.

Species	Collection	Catalog Number	No. Paired Reads	No. UCEs Recovered	Estimated Coverage
M. fotsifotsy (north clade)	FMNH	154592	3,204,578	3,056	8.72191
M. fotsifotsy (north clade)	FMNH	156312	3,645,339	2,963	6.102467
M. fotsifotsy (north clade)	FMNH	172653	3,624,190	2,906	6.301508
M. fotsifotsy (widespread clade)	FMNH	156424	519,400	1,981	1.68932
M. fotsifotsy (widespread clade)	FMNH	156568	2,963,599	2,756	6.197674
M. fotsifotsy (widespread clade)	FMNH	161959	3,084,016	3,171	6.391167
M. fotsifotsy (widespread clade)	FMNH	166142	2,941,706	2,535	5.596413
M. fotsifotsy (widespread clade)	FMNH	167428	1,410,298	2,417	4.045455
M. fotsifotsy (widespread clade)	FMNH	170750	3,978,843	2,817	6.38018
M. soricoides	FMNH	156595	4,168,491	3,067	6.984485
M. soricoides	FMNH	159460	3,085,864	3,134	6.194842
M. soricoides	FMNH	162000	3,872,895	2,789	6.424797
M. soricoides	FMNH	166121	1,892,215	2,235	4.854354
M. soricoides	FMNH	166175	2,925,109	2,659	6.416196
M. soricoides	FMNH	166206	4,088,137	2,895	6.770588
M. soricoides	FMNH	167442	2,239,297	2,863	5.625323
M. cowani	FMNH	161928	4,292,564	2,934	6.973529
M. cowani	FMNH	166102	4,551,547	2,921	7.354839

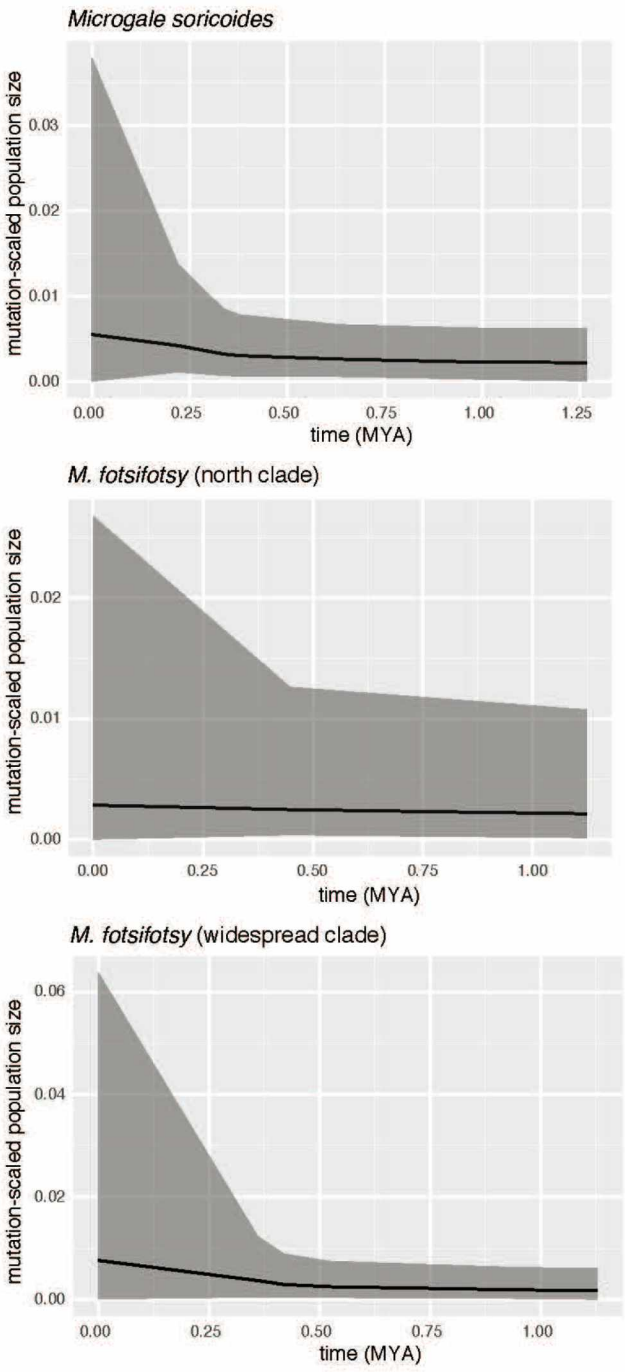
No. loci in data matrix	2,836
Total concatenated length (bp)	2,006,999
Avg locus length	708
Min locus length	191
Max locus length	1,548
Avg % missing data per locus	8.31

Supplementary Table 4.3. All morphological measurements included in this study. See Table 4.1 for measurement descriptions. Collection abbreviations are as follows: Field Museum of Natural History (FMNH), Université d'Antananarivo Département de Biologie Animale (UADBA), American Museum of Natural History (AMNH), Smithsonian Institution National Museum of Natural History (USNM), British Museum of Natural History (BMNH), University of Michigan Museum of Zoology (UMMZ). This table is available upon request, and will be made publicly available upon this chapter's acceptance in a peer-reviewed journal.

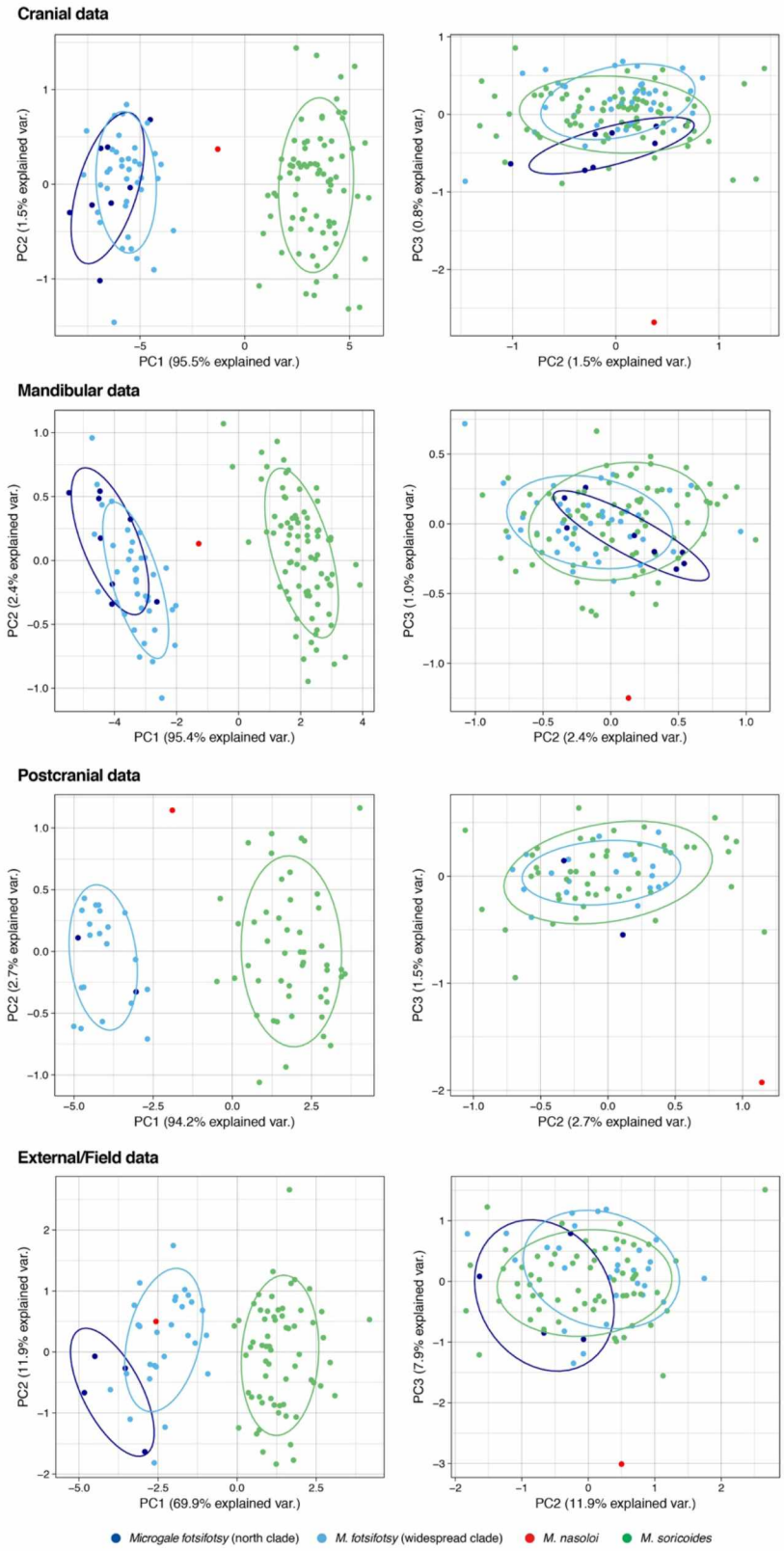
4.11 Supplementary Figures



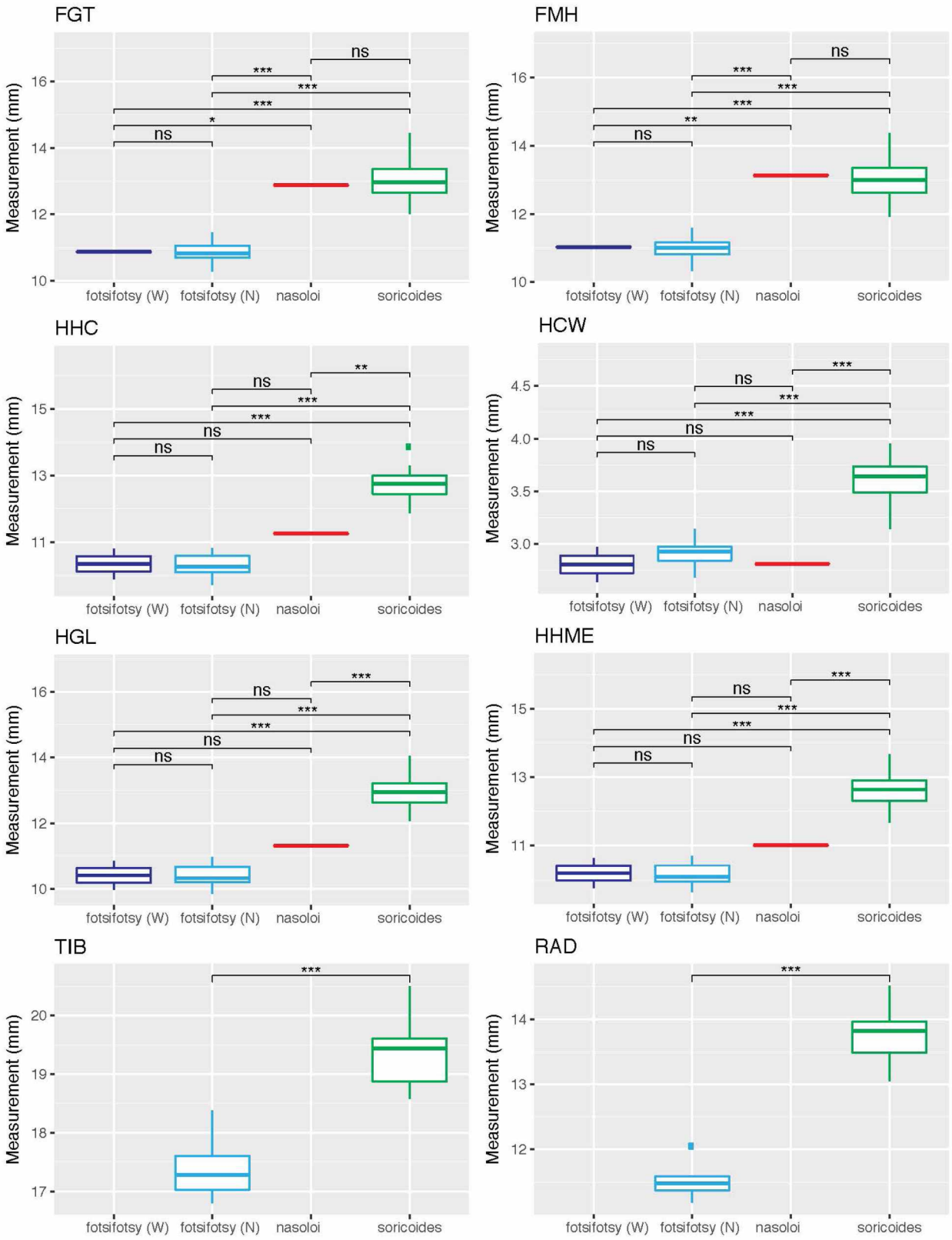
Supplementary Figure 4.1. A schematic diagram of all demographic models that were evaluated in our PHRAPL analyses. Arrows indicate the presence of bi-directional gene flow between species.



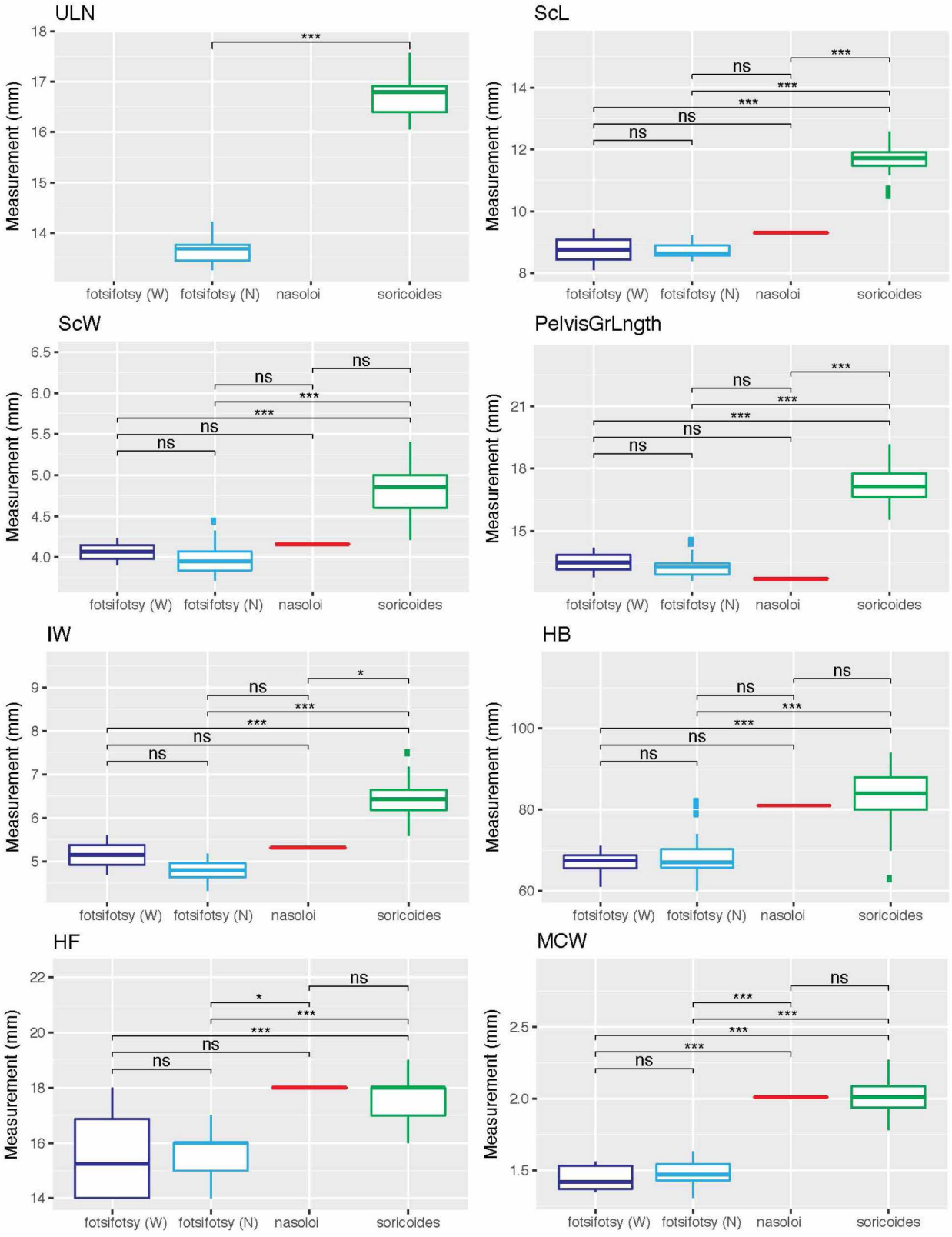
Supplementary Figure 4.2. Extended Bayesian skyline plots showing the median mutation-scaled population size through time, with the present at the origin.



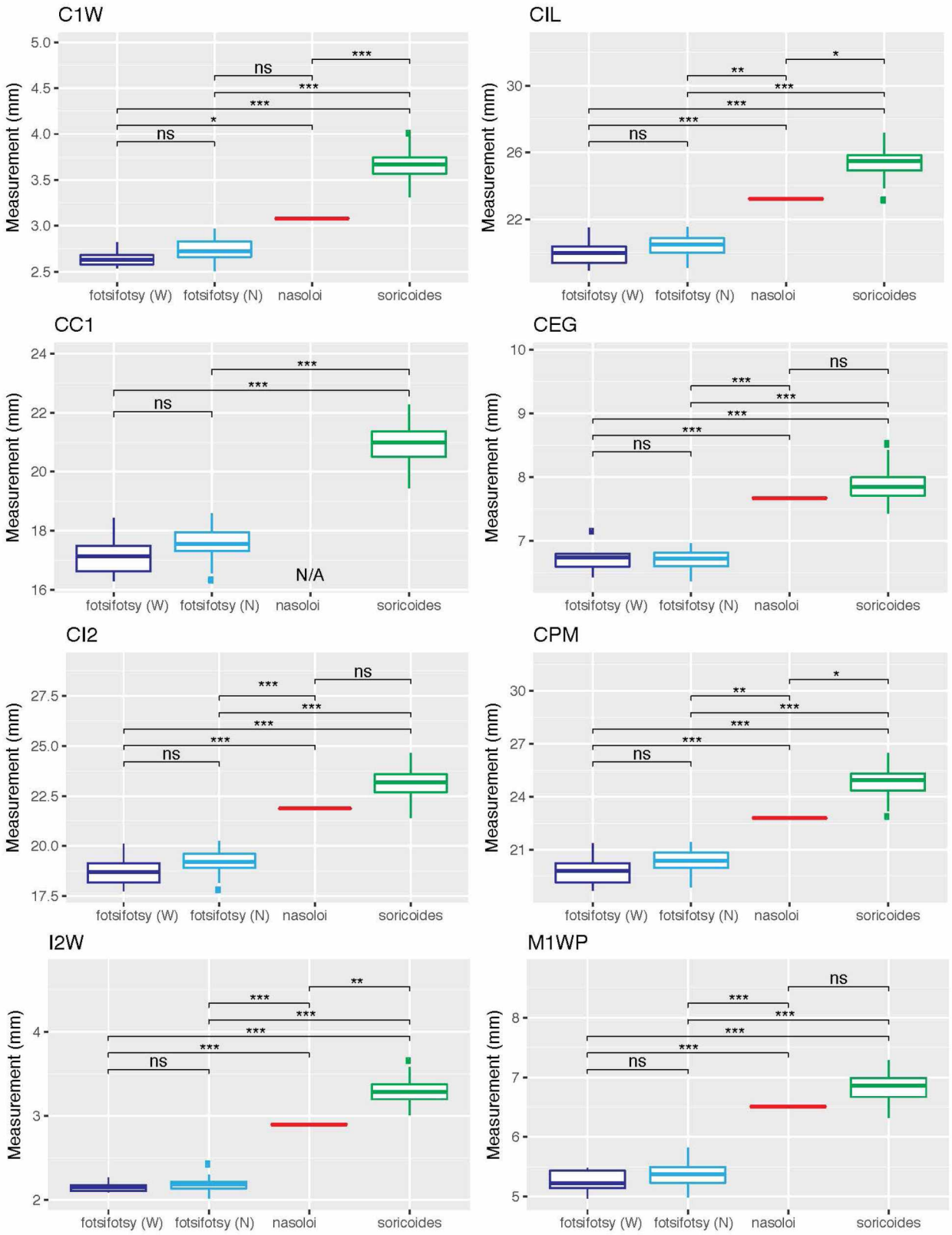
Supplementary Figure 4.3. Bivariate plots of the first two principal components (left) and the second and third principal components (right) for each morphometric module. Individual points (and 95% confidence ellipses for clusters with >2 individuals) are colored according to species, as in Fig. 4.2.



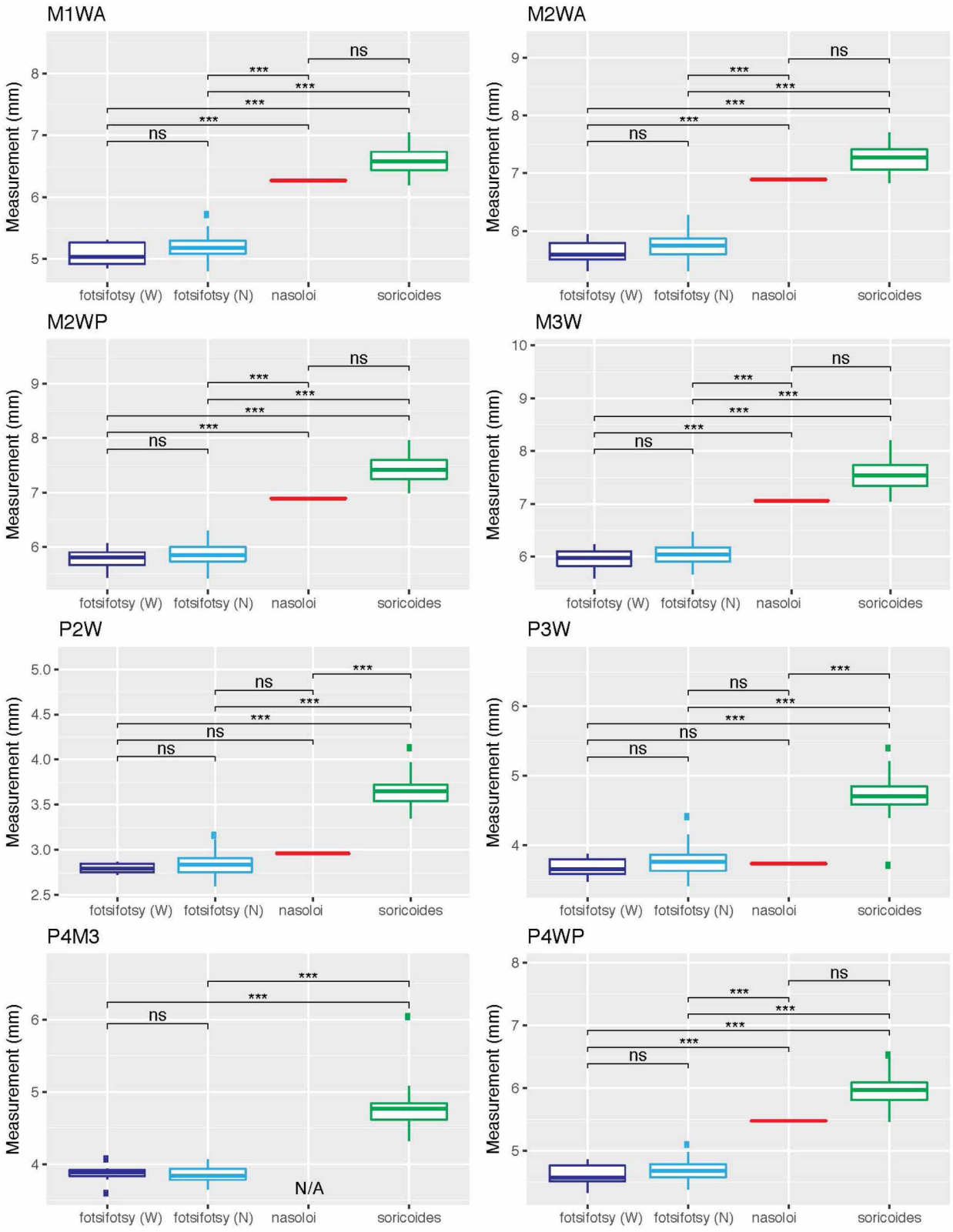
Supplementary Figure 4.4. Boxplots of all morphological measurements evaluated in this study. Brackets above each boxplot indicate significance values from pairwise ANOVAs. Boxes are colored according to species, as in Fig. 4.2. (N) and (W) refer to the north and widespread clades, respectively.



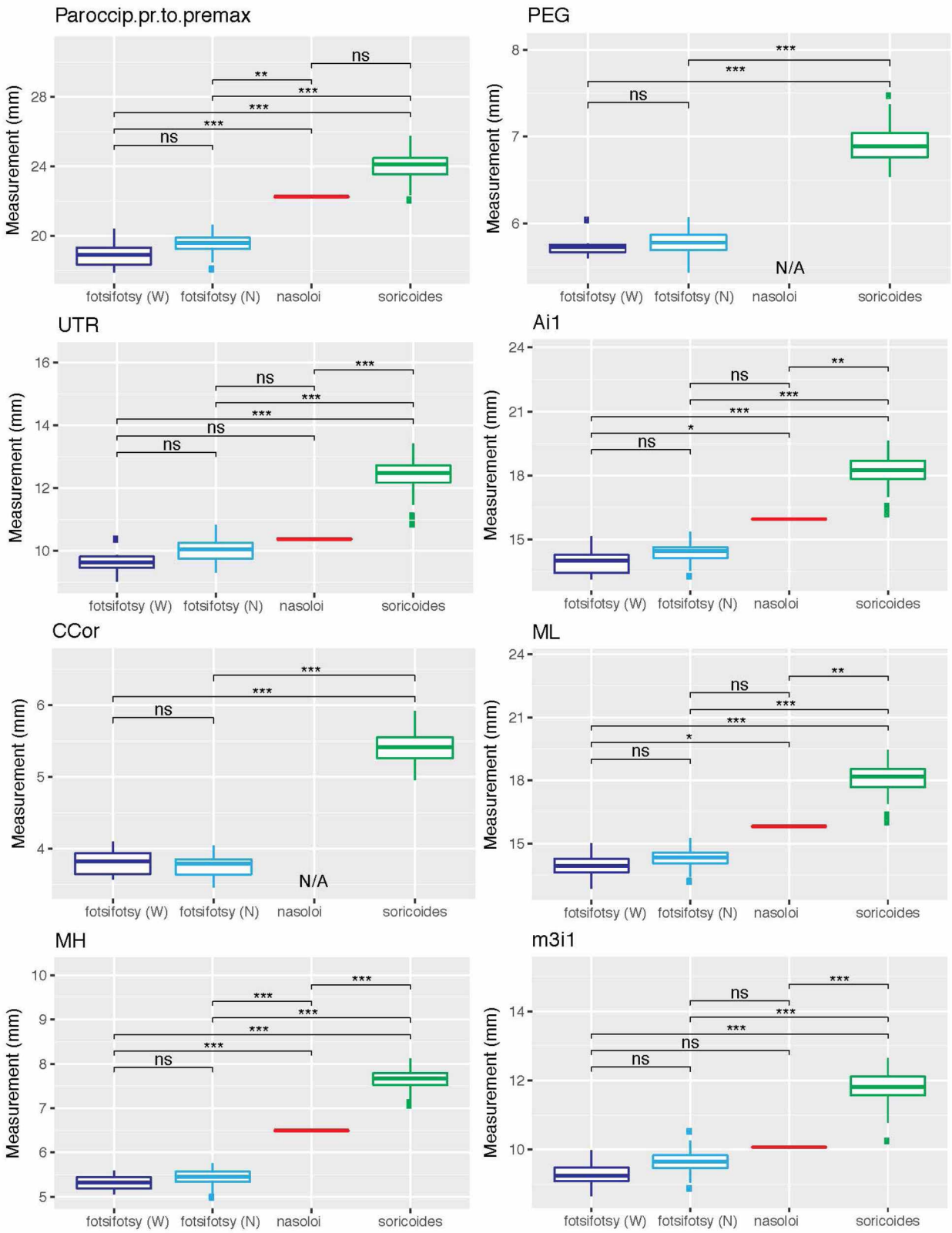
Supplementary Figure 4.4 (continued).



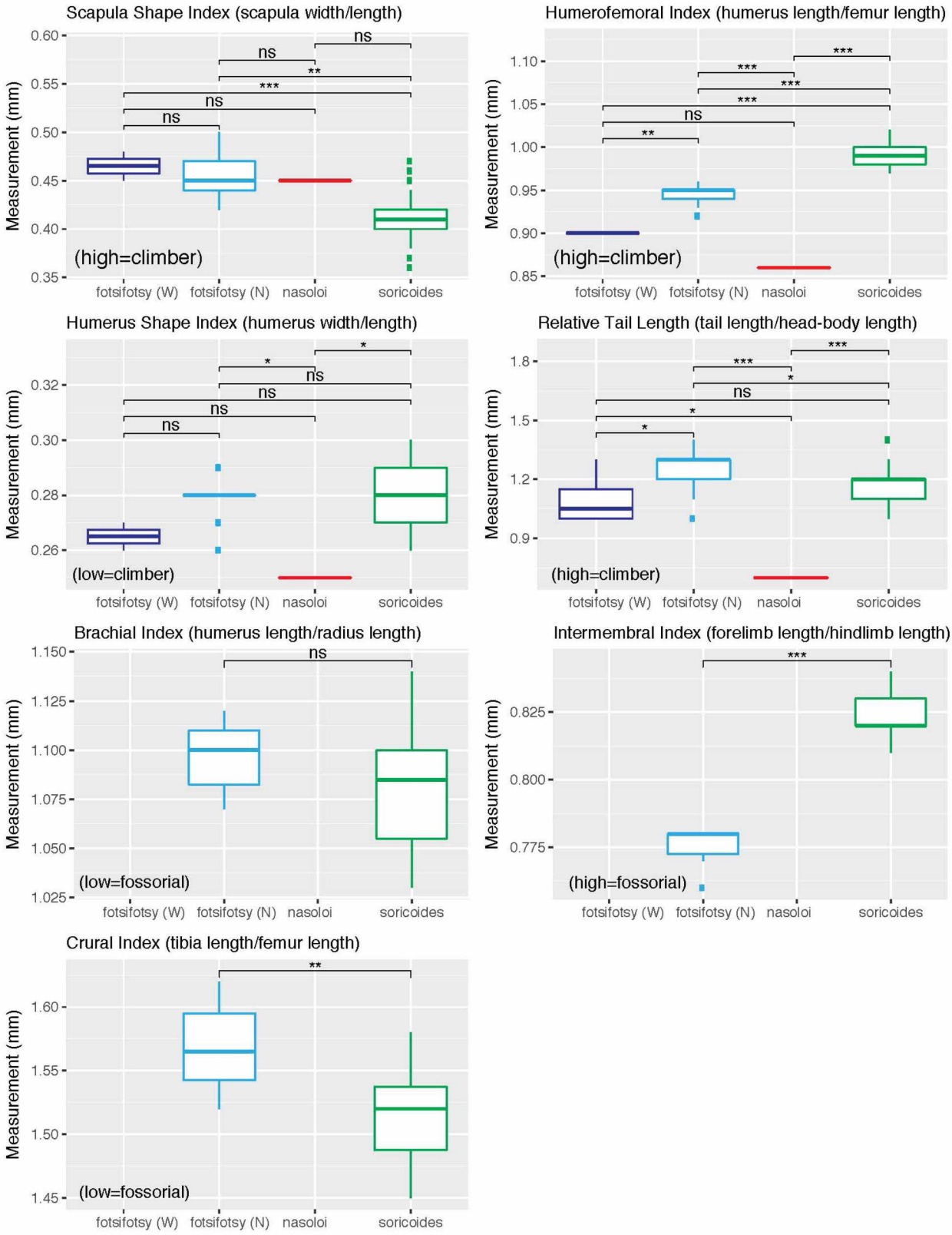
Supplementary Figure 4.4 (continued).



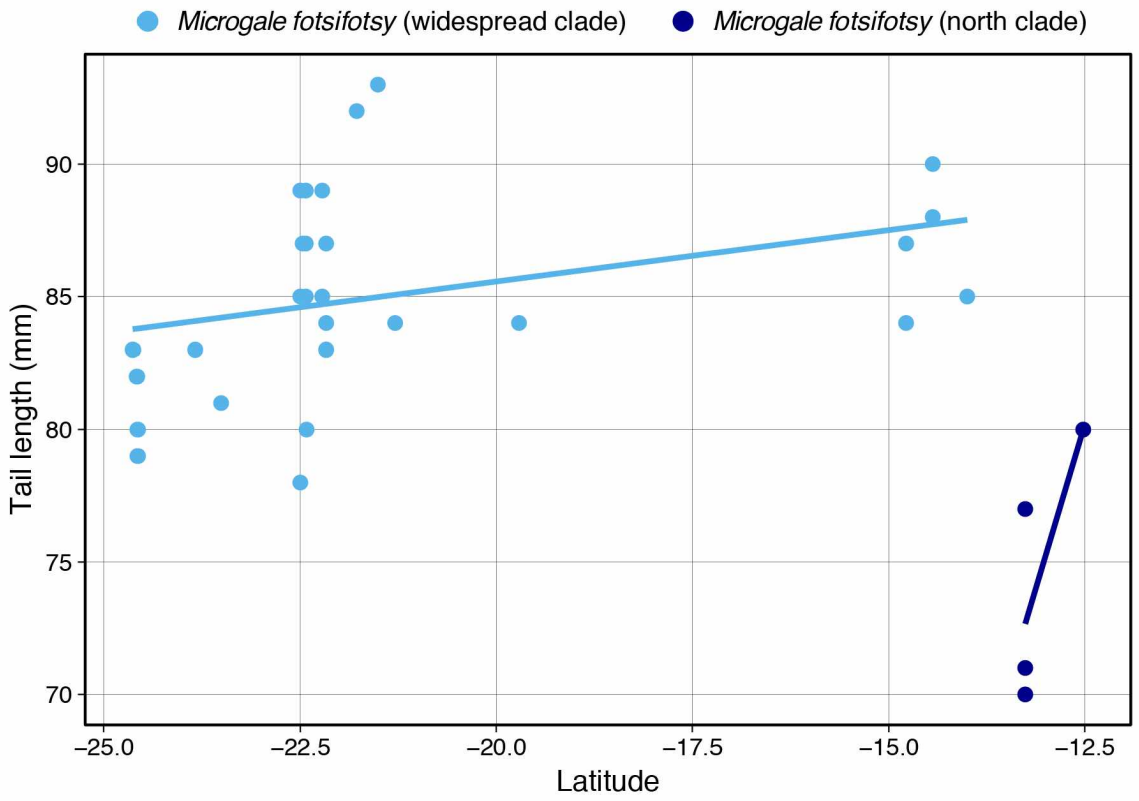
Supplementary Figure 4.4 (continued).



Supplementary Figure 4.4 (continued).



Supplementary Figure 4.4 (continued).



Supplementary Figure 4.5. Linear regression of tail length versus latitude for *Microgale fotsifotsy*. Note the sharp break in clinal variation between the widespread and north clades.

General Conclusion

In this dissertation, I explored the evolutionary history of tenrecs at the levels of family (Chapter 1), species (Chapters 2 and 4), and population (Chapters 2 and 3). In Chapter 1, I generated the first phylogeny of tenrecs to include all recognized species, finally resolving long-standing confusion about several branches of the tenrec tree. Doing so enabled me to revise the taxonomy of Tenrecidae to maximize consistency with phylogeny, evolutionary history, and ecomorphological distinctiveness; namely, I placed the web-footed tenrec in the genus *Microgale* and recognized the otter shrews of continental Africa as a separate family (Potamogalidae). Using fossil data, I also generated a time-calibrated phylogeny, revealing that the tenrec ancestor colonized Madagascar 30-56 million years ago. After colonizing the island, tenrecs underwent an extensive radiation, with my analyses showing that the greatest rates of speciation occurred in humid habitats.

In Chapter 2, to better understand the patterns and processes of diversification in Madagascar's humid forest, I focused my attention on a taxon that is restricted to humid-forest habitats, the mole tenrecs (*Oryzorictes* spp.). I found that one currently recognized species (*O. hova*) is actually a complex of at least three species. Despite there being only subtle differences in morphology, my integration of morphological, molecular, and geographic data show decisively that these species have been evolving independently for over one million years (although I cannot yet formally describe new species due to problems associated with the holotypes). The three species are latitudinally distributed and likely associated with Madagascar's northern, central, and southern highland regions. This pattern, in addition to estimates of population size through time, lends support to the existing hypothesis that Madagascar's humid forests have undergone cycles of expansion and contraction, which may be a driving force behind Madagascar's exceptional levels of species richness and endemism.

As part of my species delimitation procedure in Chapter 2, I also assessed whether two popular species validation programs (BP&P and iBPP) are sensitive to type I error, (i.e., false positives). After performing standard analyses with both methods, I replicated the analyses using randomized species assignments, which revealed that these methods are indeed susceptible to false positives. This result is an important cautionary tale to the growing number of researchers employing these methods.

The latitudinal phylogeographic pattern I identified in Chapter 2 was similar to patterns previously recognized in many other taxa from Madagascar's eastern humid forests; however, no previous study had ever looked across multiple species to explicitly test whether this pattern was related to the distribution of Madagascar's highland regions. In Chapter 3, I used comparative phylogeographic methods on a dataset of mitochondrial DNA sequence data from 13 species of tenrecs, 7 species of nesomyine rodents, and 5 species of reptiles to determine whether these phylogeographic patterns are

congruent across multiple species, and whether they can be explained by the distribution of Madagascar's high-elevation regions. Across all of the evaluated species, I identified a remarkably congruent phylogeographic structure – both spatially and temporally – that corresponds to the northern, central, and southern highlands, and the far northern Montagne d'Ambre. This work has important implications for conservation, including highlighting specific geographic regions that are critically important for generating and harboring biodiversity, and revealing several instances of cryptic endemism and incipient speciation.

Finally, in Chapter 4, I studied species limits in a clade of shrew tenrecs that includes three currently recognized species that had been suspected of occasional hybridization in the areas where their ranges overlap: *Microgale nasoloi, M. fotsifotsy,* and *M. soricoides*. Although the species are quite distinct morphologically, I found evidence of past hybridization in areas of sympatry. I also found evidence for a reproductively isolated population of *M. fotsifotsy* that occurs only in the far north of the island – the only area where its range does not overlap with *M. soricoides* – adding support to my findings from Chapter 3 that the far north harbors locally endemic cryptic species. Genetic and geographic evidence suggests that the two clades of *M. fotsifotsy* are completely reproductively isolated; this fact, in combination with some significant differences in morphological measurements, lead me to conclude that these are two different species.

In this dissertation I clarified the evolutionary relationships among tenrecs and some of the forces driving their diversification. Importantly, I demonstrated that the current number of recognized species is almost certainly an underestimate, and future researchers should continue to document and delimit tenrec species diversity before habitats disappear. I also illuminated how Madagascar's flora and fauna likely became so diverse: I revealed that speciation is greatest in the humid forest, that populations in that region have undergone population expansion and contraction, and that this has likely been related to the presence of high-elevation refugia during past climatic oscillations. These insights were made possible thanks to more than a century of biodiversity collection and preservation efforts, and I expect that the specimens used in this research will continue to provide important insights for other researchers in perpetuity.